

Familiebasert behandling av alvorlige atferdsproblemer blant barn og ungdom

Evaluering og implementering av evidensbaserte
behandlingsprogrammer i Norge

Terje Ogden

Avhandling for graden doctor philosophiae (dr. philos)

Universitetet i Bergen 2010

Forord

Forskningsprosjektene som rapporteres i denne avhandlingen hadde ikke vært mulige uten bidrag fra mange personer og miljøer. Det gjelder for det første engasjerte PMTO og MST terapeuter som har utført behandlingene som utgjør grunnlaget for studiene. Videre gjelder mine kolleger i utviklingsavdelingene for barn og ungdom som har hatt ansvar for opplæring og veiledning av terapeutene og som har støttet forskningsprosjektet i alle dets faser. En stor takk til avdelingsdirektørene Eliabeth Askeland og Bernadette Christensen og deres kolleger (Anett Apeland, Terje Christiansen, Mona Duckert, Tori Mauseth og Roar Solholm). En særlig takk går til min kollega og medforfatter Kristine Amlund-Hagen for inspirerende og produktivt samarbeid om hele fire av de fem artiklene som inngår i avhandlingen. En takk også til Colleen Halliday-Boykins som bidro i den første MST studien. Jeg vil også takke alle mine kolleger i forskningsavdelingen ved Atferdssenteret for gode råd og kommentarer. Ikke minst har logistikk teamet hatt viktige oppgaver i forbindelse med kontakt og oppfølging av respondentene, blant andre John Kjøbli, Asgeir Olseth og i tillegg til disse Carl Ivar Holmen, Eyvind Elgesem og Torunn Ogden. Kristin Nordahl har vært en viktig støttespiller gjennom å lede logistikk teamet og ikke minst gjennom å organisere kodingen av observasjonsdata. Arbeidet som undersøkelsene bygger på har blitt finansiert av flere departementer, men først og fremst av Barne- og Likestillingsdepartementet, der Haktor Helland har vært en interessert og engasjert støttespiller. Vi har i prosessen også hatt god støtte fra programutviklerne Scott Henggeler, Marion Forgatch og Gerald Patterson som har delt sine erfaringer med oss og kommet med viktige innspill og forslag i prosessen. Til slutt vil jeg også takke familiene som bidro i prosjektet gjennom å stille opp til intervjuer og video-observasjoner og som på denne måten delte sine erfaringer og vurderinger med oss og andre som er opptatt av å bedre tilbudet til barn og unge med alvorlige atferdsproblemer.

Moss/Oslo mai 2010

Terje Ogden

Innholdsfortegnelse

Forord.....	s. 2.
Sammenheng s. 4.	
Liste over artikler	s. 6.
1. Utviklingen av nye tiltak i arbeidet med alvorlige atferdsproblemer blant barn og unge i Norge	s. 7.
2. Evidensbasert policy og utviklingen av evidensbaserte behandlingsprogrammer.	s. 21
3. Evidensbaserte behandlingsprogrammer – programfaktorer, fellesfaktorer.. og klinisk skjønn.....	s 41.
4. Randomiserte kontrollerte evalueringsstudier – det eksperimentelle designet.....	s. 68.
5. Parent Management Training, Oregon og Multisystemisk Terapi – programkennetegn, og forskning	s. 90.
6. Presentasjon av undersøkelsene	s. 114.
7. Avsluttende drøfting.....	s. 135.
8. Litteraturreferanser.....	s. 146.
9. Appendiks.....	s. 171.

Sammendrag

Artikkel 1 om den norske PMTO (Parent Management Training Oregon modellen) studien viste at foreldrebehandling kan ha en signifikant positiv effekt på barns problematferd målt ved avslutningen av behandlingen i et klinisk utvalg. Positive hovedeffekter ble målt for foreldrevurderinger av barnas samlede problematferd og eksternaliserte problemer, men også for foreldrenes daglige rapporteringer over tre dager. PMTO foreldrene ga også mer positive vurderinger av behandlingstilbudet enn foreldrene i sammenligningsgruppen. I følge lærervurderingene forbedret barna i PMTO gruppen sine sosiale ferdigheter signifikant i løpet av behandlingen sammenlignet med de andre barna, noe som tyder på generaliseringseffekter fra hjem til skole. Resultatene underbygde at grensesetting var den av foreldreferdighetene som var mest avgjørende for de signifikant positive endringene i PMTO familienes favør. Blant interaksjonseffektene var alder den tydeligste moderatoren som viste at for barn under 8 år var resultatene signifikant mer positive for PMTO enn for kontrollbetingelsene på flere variabler, både i foreldre- og lærervurderinger.

I artikkel 2 rapporteres en randomisert kontrollert studie som viste at Multisystemisk terapi var mer virksomt enn tjenestetilbudet som vanligvis ble gitt til norske ungdommer med alvorlige atferdsproblemer som ble henvist til Barneverntjenesten. Evalueringen umiddelbart etter avslutning av behandlingen og oppfølgingsstudien 2 år etter inntak, viste begge at MST ungdommene hadde en mer positiv utvikling enn de som hadde fått andre tilbud og annen oppfølging fra Barneverntjenesten. Det er særlig grunn til å framheve at MST i høyere grad enn andre tiltak hadde forebygget at ungdom ble plassert utenfor hjemmet.

I artikkel 3, oppfølgingsstudien, rapporterte foreldre og lærere i MST gruppen også om signifikant mindre eksternalisert problematferd og ungdommene om færre lov- og regelbrudd enn informantene i kontrollgruppen. Når resultatene var inkonsistente for ungdommenes selvrappotering kan det tyde på at MST hadde en målrettet effekt på problematferden som behandlingen fokuserte på, mens effekten var mindre på andre selvrappoterte vansker. Samtidig viste det seg noe uventet at foreldrene og lærerne i MST gruppen rapporterte mindre internalisert problematferd enn foreldre og lærere i sammenligningsgruppen.

I artikkel 4 om kjønnsforskjeller i MST viste det seg at selv om jentene hadde en noe annen problemprofil enn guttene ved inntak, så endret de seg like mye i løpet av behandlingen. Jentene som ble henvist til MST var sammenlignet med guttene like

problembelastet med hensyn til utagerende atferd og rusmisbruk, de hadde like mange tilleggsp problemer, men de startet oftere sin problemkarrière på et senere tidspunkt. Selv om gjennomsnittsalderen for henvisning til MST var den samme for begge kjønn (15 år), var det klare indikasjoner på at flere gutter enn jenter hadde utviklet problemer før 10 års alderen. Undersøkelsen viste at jentene endret seg like mye som guttene. Men selv om MST var en virksom behandling for de fleste ungdommene i utvalget, så kunne ikke resultatene underbygge at det var en kausal sammenheng mellom behandling og resultater.

Artikkel 5 er en kasusstudie av komponentene som inngår i prosessen der en utvikler, implementerer og evaluerer evidensbaserte behandlingsprogrammer. Artikkelen beskriver en modell basert på de norske erfaringene med formidling fra programutviklerne, adopsjon av brukerorganisasjonene, implementering som foregår i et avveiningsforhold mellom høy grad av behandlingsintegritet og behovet for lokal tilpasning og til slutt, evaluering av resultatene. Implementeringen av MST og PMTO ble organisert som en kombinasjon av ”top-down” og ”bottom-up” strategier slik at programmene ble adoptert basert på råd fra en ekspertgruppe, finansiert av departementer og den fylkeskommunale forvaltningen og implementert i et samarbeid mellom et nasjonalt senter, barneverntjenesten og psykisk helsevern. Forhold som virket fremmende på implementeringen av PMTO og MST i Norge var en generell bekymring for underkapasitet og manglende kvalitet på behandlingstilbudet til barn med alvorlige atferdsproblemer, en generell etterspørsel etter og interesse for evidensbaserte behandlingsmetoder blant praktikere, og en vilje i forvaltningen til å finansiere og støtte implementering og forskning om behandlingsmetodene. Positive brukererfaringer og tilbakemeldinger fra familier som hadde gjennomgått behandlingen var også medvirkende til at et økende antall familier ble henvist eller tok kontakt for å få disse behandlingstilbudene.

Undersøkelsene som inngår i denne avhandlingen underbygger at familien er en viktig bidragsyter i arbeidet med å endre alvorlige atferdsproblemer blant barn og unge. Studiene viser at de foresatte er viktige aktører både i tiltak som fokuserer på tidlig intervensjon og tiltak som retter seg mot ungdom med alvorlige og sammensatte atferdsproblemer. Undersøkelsene viser også at evidensbaserte behandlingsprogrammer kan bidra til at foreldrene blir mer kompetente og trygge slik at de møter sine barn på måter som fører til færre konflikter, et bedre samhold i familien og en reduksjon av utagerende problematferd. Det er også et viktig moment at programmene bygger på prinsippet om familiebevaring og behandler problemene der de oppstår, det vil si i barnas og ungdommenes hjemmemiljø.

Artikler

Artikkel 1.

Ogden, T. & Amlund-Hagen, K. (2008). Treatment effectiveness of Parent Management Training in Norway: A randomized controlled trial of children with conduct problems. *Journal of Consulting and Clinical Psychology*, 76, 607-621.

Artikkel 2.

Ogden, T. & Halliday-Boykins, C.A. (2004). Multisystemic Treatment of Antisocial Adolescents in Norway: Replication of Clinical Outcomes Outside of the US. *Child and Adolescent Mental Health*, 9, 77-83.

Artikkel 3.

Ogden, T., & Amlund-Hagen, K. (2006). Multisystemic therapy of serious behaviour problems in youth: Sustainability of therapy effectiveness two years after intake. *Child and Adolescent Mental Health*, 11, 142-149.

Artikkel 4.

Ogden, T. & Amlund-Hagen, K. (2009). What works for whom? Gender differences in intake characteristics and treatment benefit of Multisystemic Therapy. *Journal of Adolescence*, 32, 1425-1435.

Artikkel 5.

Ogden, T., Amlund-Hagen, K., Askeland, E., & Christensen, B. (2009). Implementing and evaluating evidence-based treatments of conduct problems in children and youth in Norway. *Research on Social Work Practice*, 19, 582-591.

The published papers are reprinted with permission from the publishers. All rights reserved.

1. Utviklingen av nye tiltak i arbeidet med alvorlige atferdsproblemer blant barn og unge i Norge

Alvorlige atferdsproblemer hos barn og unge handler om store og vedvarende avvik fra normer og regler for atferd, om alvorlig krenkelse av andres rettigheter eller lovbrudd. Hos barn dreier det seg om aggresjon samt norm- og regelbrytende atferd som når de er vedvarende ulydige, får raseriutbrudd, har et heftig sinne, er uvanlig bråkete eller går løs på andre fysisk (Achenbach, 1991). Hos ungdom handler det i tillegg om vold, lovbrudd, rusmisbruk, rømming og problemer i skolen, for eksempel skulk. Alvorlige atferdsproblemer kan være stabile gjennom barndom og ungdomsalder og representerer en betydelig risikofaktor for for framtidige vansker som avbrutt skolegang, misbruk, kriminalitet, psykiske helseproblemer, dårlig tilpasning til arbeidslivet, familieproblemer samt høyere sykkelighet og dødelighet (Kjellsberg & Dahl, 1998; Walker, Ramsey & Gresham, 2004). Det er derfor all grunn til å fokusere på hva som kan gjøres for å stoppe eller begrense problemkarrierer i oppveksten. Intervensjonsforskningen peker på viktigheten av tidlig identifikasjon og intervensjon, og at virksomme tiltak for eldre barn og ungdom må være intensive og omfattende (Loeber & Farrington, 1998; Kazdin, 1997).

Synet på nytten av psykososial behandling av alvorlige atferdsproblemer har endret seg i løpet av de siste 50 årene. I tiårene etter den andre verdenskrigen rådet en positiv forventning til at det var mulig å redusere alvorlige atferdsproblemer gjennom miljøtiltak og miljøskifte blant annet gjennom plasseringer i fosterhjem, spesialskoler og institusjoner. Holdningen gjennom 1950-60 årene kan beskrives med uttrykket ”alt virker”. Man tenkte seg at alle tiltak som ble iverksatt for vanskelige barn og unge, fra plasseringstiltak til sysselsettings- og fritidstiltak var et ubetinget gode, selv om lite ble evaluert (Innstilling, 1968). Internasjonal forskning på 1970 tallet modererte denne optimismen med påstander om at det var praktisk talt nytteløst å behandle alvorlige atferdsproblemer. Selv om slagordet ”ingenting nytter” formulert av Martinson (1974) viste seg å være *for* pessimistisk, så bidro det til å redusere tiltroen til nytten av psykososial behandling og det ble satset mer på strafferettslige reaksjoner og plasseringstiltak (Lösel, 1993). Kvaliteten på forsøk, forskning og utviklingsarbeid økte også i 1980-90 årene og dokumenterte at behandlingstiltak kunne gi positive og varige resultater i arbeidet med de vanskeligste (Loeber & Farrington, 1998; Forgatch, 1991; Borduin mfl., 1995). Ikke minst bidro utviklingen av strukturerte behandlingsprogrammer som bygde på og ble evaluert gjennom forskning, til økt optimisme

(Biglan & Ogden, 2008). Systematiske forskningsoversikter viste at de mest virksomme tiltakene besto av åpen behandling i familie og nærmiljø, eventuelt i kombinasjon med plassering i institusjon og fosterhjem (Kazdin, 1997; Lösel, 1993). I det siste tiåret har utviklingen av empiriske modellprogram for forebygging og behandling av alvorlige atferdsproblemer fortsatt (se blant annet oversikter hos Carpelan mfl., 2008; APA, 2008; Lipsey & Cullen, 2007; Piquero, Farrington, Welsh, Tremblay & Jennings, 2008). Med slagordet "what works" la behandlingsprogrammene stor vekt på evaluering av kliniske behandlingsresultater og utfordret dermed den sterke prosessorienteringen som hadde dominert atferdsfeltet. Mens en tidligere la stor vekt på behandlingens teoretiske grunnlag, innhold og kvaliteter ble en nå mer opptatt av resultatevaluering gjennom kontrollerte evalueringsstudier. Det førte til en økt bevissthet om betydningen av å dokumentere og å synliggjøre metodikk og arbeidsmåter i behandlingen av alvorlige atferdsproblemer. Primærstudier av programmenes behandlingseffekter ble etter hvert oppsummert i systematiske kunnskapsoversikter og meta-analyser (Aos, Lieb, Mayfield, Miller, & Pennucci, 2004; Drake, Aos & Miller, 2009; Kaminski, Valle, Filene & Boyle, 2008).

Barne og familiedepartementet la i 1997 fram en «Handlingsplan for barn og unge med alvorlige atferdsproblemer» (Rundskriv Q-12/97). Hensikten var å sikre et bedre og mer helhetlig hjelpetilbud både til barn og unge som allerede hadde utviklet og de som sto i fare for å utvikle alvorlige problemer. For å styrke innsatsen ble det blant annet foreslått å: 1) tydeliggjøre ansvarsforholdene i forvaltningen og mellom de ulike forvaltningsnivåene, 2) bygge ut varierte fylkeskommunale tiltak og institusjonsplasser, 3) utvikle en fylkeskommunal konsultasjonstjeneste og 4) styrke forskning og kompetanseutvikling. Det kan være grunn til å merke seg at handlingsplanen la vekt på å styrke rammebetingelsene for et bedre og bredere tjenestetilbud til målgruppen basert på politiske vurderinger av situasjonen i praksisfeltet. Men planen inneholdt lite om hvilke tiltak og tjenester som skulle prioriteres og hadde få referanser til forskningen på feltet. Men høsten 1997 arrangerte Barne- og familiedepartementet og styret for Norges Forskningsråds program for Mental helse en ekspertkonferanse om «barn og unge som er særlig vanskelige og utagerende». Til konferansen ble ledende forskere på feltet, primært fra USA og Canada, invitert til å presentere sin forskning om intervensjoner som hadde gitt lovende resultater i behandlingen av utsatte barn og unge.

En ekspertgruppe ble etablert i forbindelse med konferansen og utarbeidet senere en rapport med tittelen «Barn og unge med alvorlige atferdsvansker, hva kan nyere viten fortelle oss og hva slags hjelp trenger de?» (Zeiner mfl., 1998). Rapporten inneholdt anbefalinger om videre metode- og tiltaksutvikling og var preget av den internasjonale forskningsstatusen på feltet. Blant annet framhevet den betydningen av tidlig identifikasjon og intervensjon og satsing på familie- og nærmiljøbaserte tiltak. Den påpekte også de potensielle negative effektene av å behandle alvorlige atferdsproblemer i gruppetiltak (Dishion, McCord & Poulin, 1999). I rapporten anbefalte også ekspertgruppen at Norges Forskningsråd i samarbeid med berørte departementer burde sette i gang utprøving og evaluering av forskningsbaserte behandlingsprogram i randomiserte kontrollerte studier (Zeiner mfl., 1998). Forskningsrådets engasjement i dette arbeidet ble av kort varighet, og ingen av rådets styrer tok initiativ til finansiering av intervensjonsforskning på atferdsfeltet. Det var departementene, og særlig Barne- og familiedepartementet (BFD) som førte prosessen videre.

Som en oppfølging av sitt rundskriv Q-12/97 og i tråd med anbefalingene fra ekspertkonferansen finansierte BFD en forskerstilling ved Universitetet i Oslo. Stillingen var begynnelsen til en forskningsenhet som ble etablert ved Psykologisk Institutt, Senter for klinisk psykologisk forskning fra 1998. Ved siden av å initiere egen forskning om behandling av atferdsvansker blant barn og unge, skulle forskeren drive opplæring av studenter i behandlingsarbeid overfor denne målgruppen. Forskningsenheten skulle også ha ansvar for å formidle ny kunnskap til de fylkeskommunale barne- og ungdomstjenestene i form av kursvirksomhet og foredrag. Senere skulle det vise seg at ved siden av forskningen kom arbeidet i stor grad til å dreie seg om tiltaksutvikling og kompetansebygging i praksisfeltet med utgangspunkt i strukturerte behandlingsprogrammer.

I forbindelse med ekspertkonferansen høsten 1997 ble det også avholdt to dagskurs i klinisk forskningsmetode for interesserte forskere og andre fagpersoner. Den ene dagen hadde Gerald Patterson og Marion Forgatch fra Oregon Social Learning Center (OSLC) ansvaret for og den andre dagen foreleste Scott Henggeler og Melisa Rowland fra Family Services Research Center, Medical University of South Carolina (MUSC). I løpet av kursdagene ble det etablert kontakter som resulterte i norske studiebesøk i Eugene Oregon og Charleston, South Carolina. Representanter fra BFD, Universitetet i Oslo og flere praksismiljøer deltok på disse studiebesøkene, og det ble det gjort avtaler om å implementere Oregon modellen av Parent Management Training (PMTO) (Ogden, 1998) og Multisystemisk Terapi (MST) (Ogden, 1999) i Norge. Med støtte først og fremst fra BFD startet enheten ved Psykologisk Institutt et forsknings- og utviklingsarbeid som fikk tittelen "Atferdsprosjektet". Til prosjektet

ble det i løpet av 1998-1999 knyttet en faggruppe på 6 personer som fikk i oppgave å implementere og evaluere programmene på nasjonal basis. Barne- og familiedepartementet (BFD) tok sammen med Sosial og helsedepartementet (SHD) initiativ til et møte med fylkeshelsesjefene der de tilbød fylkene et samarbeid om implementering av metodene. Av 19 fylker tok 17 i mot tilbudet om å opprette MST team i den fylkeskommunale Barneverntjenesten, og PMTO som skulle etableres som et poliklinisk tilbud i Barneverntjenesten og Barne- og ungdomspsykiatrien. MST-services tilknyttet Family Services Research Center ved MUSC fikk ansvaret for opplæring og konsultasjon av de norske MST teamene. Videre påtok Marion Forgatch, Gerald Patterson og deres kolleger ved OSLC seg oppgaven med å lære opp aktuelle kandidater i PMTO. Begge programmene ble introdusert i Norge høsten 1999.

Denne avhandlingen beskriver arbeidet med å og evaluere og implementere MST (Henggeler, Schoenwald, Borudin, Rowland & Cunningham, 1998, 2009; Ogden, Christensen, Sheidow & Holth, 2008) og PMTO (Ogden, Forgatch, Askeland, Patterson & Bullock, 2005) i alle helseregionene i Norge. Resten av dette introduksjonskapitlet inneholder en begrepsanalyse av alvorlige atferdsproblemer hos barn og unge og gir deretter en kort oversikt over tjeneste- og reaksjonssystemet i forhold til slike problemer. I kapittel 2 omtales forholdet mellom evidensbasert policy og praksis med vekt på spørsmålet om hvor forpliktende forholdet mellom forskning og praksis skal være. Videre beskrives en begrepsmodell for utviklingen og implementeringen av evidensbaserte programmer samt kriterier for evidens. I det neste kapitlet, kapittel 3, drøftes evidensbaserte behandlingsprogrammer mer inngående med vekt på temaer som behandlingsprotokoller, behandlingsintegritet, komorbiditet og behandlingstidens lengde. Kapitlet tar også opp forholdet mellom programbasert behandling og behandling basert på klinisk skjønn. Kapittel 4 presenterer Oregon modellen for Parent Management Training (PMTO) og Multisystemisk terapi (MST) med vekt på deres teorigrunnlag, kjernekomponenter og forskningsresultater. Kapittel 5 gjennomgår det eksperimentelle designet og premisser for rapportering av randomiserte kontrollerte forsøk. Kapittel 6 presenterer så undersøkelsene som inngår i avhandlingen mens kapittel 7 inneholder en avsluttende diskusjon.

Alvorlige atferdsproblemer

Når atferdsproblemer omtales som *alvorlige* er det for å skille den fram vanlige og forbigående oppdragelsesproblemer eller vanlige foreldres problemer med vanlige barn. Brukt i faglig sammenheng kan alvorlig for det første vise til alvorlige enkelthandlinger som lovbrudd, vold, brannstifting og overgrep. For det andre kan «alvorlig» vise til en kombinasjon av handlinger som i sum utgjør et alvorlig problem. Diagnosen «atferdsforstyrrelse» er et eksempel på det. For det tredje kan 'alvorlig' signalisere at problemene er stabile over tid og at de er gjennomgående (eng. pervasive) slik at de kommer til uttrykk i flere miljøer som hjemme, på skolen og i fritiden.

Atferdsproblemer er sosialt definerte, og oppfatningen av dem kan derfor skifte over tid avhengig av hva som betraktes som lovbrudd eller alvorlige brudd på normer og regler. Det er blant annet vanskelig å avgjøre hvor alvorlige, omfattende eller stabile problemene skal være før de utløser samfunnsmessige reaksjoner i form av straff, plassering, behandling eller andre tiltak. Derfor knytter det seg også usikkerhet til prevalenstillene i Norge siden det ikke foreligger representative nasjonale undersøkelser på området. I en norsk studie rapporterte mødre til barn mellom halvannet og fire år at 3-4 prosent ofte var svært aggressive og vanskelige å oppdra, mens mer enn 70% var slik av og til (Mathiesen & Sanson, 2000). Tre studier fra Bergen belyser forekomsten av antisosial atferd blant ungdom: Kaufmann (1988) fant at cirka 5% av 7. klassingene sto for det meste av den alvorlige antisosiale atferden, mens Bendixen og Olweus (1999) fant at cirka 5% av guttene i alderen 13-14 år hadde blitt arrestert av politiet de siste 5 månedene. Den tredje studien, "Barn i Bergen" (Heiervang mfl., 2007) viste at 3,2% av 8-10 åringene kvalifiserte for diagnosen atferdsforstyrrelse. Andre studier viser at atferdsproblemer i Norge som i andre land forekommer hyppigere i bykommuner enn i distriktskommuner (Lavik, 1976; Wickstrøm, Skogen & Øia, 1996). Basert på foreliggende undersøkelser ser det derfor ut til at mellom 3 og 5 prosent av norske barn og unge har alvorlige atferdsproblemer.

Blant faguttrykkene som brukes om alvorlige atferdsproblemer blant barn og unge kan det skilles mellom en utviklingspsykologisk og en medisinsk-diagnostisk retning. I et utviklingspsykologisk perspektiv kan atferdsproblemer beskrives langs en kontinuerlig dimensjon som grader av antisosial atferd. I en medisinsk-diagnostisk tilnærming vil en ut fra en kategoriell tilnærming skille mellom atferd som faller innenfor det diagnostiske området, og atferd som ikke kvalifiserer for diagnosen. Atferd som etter en kvalifisert klinisk utredning

faller innenfor diagnosekriteriene omtales som atferdsforstyrrelse (eng. conduct disorder – CD) og omfatter et avgrenset utvalg av antisosiale handlinger (Ogden, 2002).

Diagnoser representerer *en kategoritilnærming* som definerer en terskelverdi for antisosial atferd, og der alle som skårer over denne antas å være i det kliniske området og kvalifiserer for behandling. Antisosial atferd representerer den såkalte *dimensjonstilnærmingen* der barn og unges gjennomsnittsskårer plasserer dem langs et kontinuum fra trivielle til alvorlige vansker. Men også her kan skårer under eller over en viss kriteriegrense på standardiserte vurderingsskalaer for eksternalisert problematferd definerer et klinisk diagnostisk område (for eksempel skårer over 90. prosentil på Child Behavior Checklist i et normativt utvalg, Achenbach, 1991).

Atferdsforstyrrelse

Atferdsforstyrrelse er en psykiatrisk diagnose som beskriver et repeterende og vedvarende mønster av aggressiv eller utfordrende atferd (ICD-10). Det dreier seg om klare brudd på sosiale forventninger i forhold til alderen. Sinneutbrudd som kan være en normal del av treåringens utvikling eller aggressive enkelthandlinger på alle aldersnivåer, gir ikke grunnlag for diagnosen. Eksempler på atferd som kan begrunne diagnosen er ifølge ICD-10 manualen (1999): ”betydelig grad av slåssing eller tyrannisering, grushomheter overfor dyr eller andre mennesker, alvorlig ødeleggelsestrang, ildspåsettelse, stjeling, gjentatt lyving, skulking fra skole eller rømming, uvanlig hyppige eller alvorlige raserianfall, utfordrende og provoserende atferd og vedvarende alvorlig ulydighet.” (ICD-10, 1999, s. 263). En formell diagnose skal stilles av kvalifisert fagpersonnell som psykolog, lege eller psykiater, og diagnosen anbefales ikke brukt med mindre atferden har vart i minst 6 måneder.

Opposisjonell atferdsforstyrrelse blir vanligvis beskrevet hos barn som er yngre enn 9-10 år og omfatter markert trass, ulydighet, provoserende atferd, men samtidig av mangel på mer alvorlige usosiale eller aggressive handlinger (ICD-10, 1999). De viktigste indikatorene er en tendens til ofte og aktivt å trosse voksnes anmodninger eller regler, og bevisst å irritere andre. De virker sinte, bitre, eller blir lett irritert på andre som de gir skylden for sine egne feil eller problemer. Videre har de ofte en lav frustrasjonstoleranse og mister lett besinnelsen, de kan virke provoserende, starter konfrontasjoner og er generelt uhøflige, lite samarbeidsvillige og i opposisjon til autoriteter. Opposisjonell atferdsforstyrrelse har mange likhetstrekk med atferdsforstyrrelse (CD), men mangler kjennetegn på fysisk aggressiv og utagerende atferd. Skillet kan derfor sies å gå mellom destruktiv og ikke-destruktiv atferd (Frick, 1998). ODD kan være en forløper for atferdsforstyrrelse, den kan forekomme samtidig, eller den kan

ansees som en mindre alvorlig variant av denne. Det er derfor uklart om det dreier seg om en kvantitativ eller kvalitativ forskjell mellom atferdsforstyrrelse og opposisjonell atferdsforstyrrelse.

Antisosial atferd

Antisosial atferd kan defineres som gjentatte brudd på sosiale normer, hjemme, på skolen og i fritiden og som er til skade for andre mennesker, for en selv eller for materielle ting som igjen får konsekvenser for andre (Kaufmann 1988). Atferden er vedvarende og forutsigbar og påvirkes ikke nevneverdig av de sosiale konsekvensene, ofte av negativ karakter, som følger den (Walker m.fl., 2004). I ungdomsalderen handler det om individer som er aggressive, begår tyveri, ran, hærverk og innbrudd, rømmer hjemmefra eller skulker skolen. Antisosial atferd har tradisjonelt blitt brukt om aggressiv og kriminell atferd hos ungdom, og voksne. Patterson (1982) har imidlertid argumentert for å omtale barns aggressive, utagerende, og hyperaktive atferd som antisosiale handlinger fordi det er det minst kontroversielle begrepet samtidig som det beskriver forløperne til alvorlig antisosial atferd. Definisjonen forutsetter at hendelser er både aversive og betingede (eng. contingent). En aversiv hendelse forekommer for eksempel når barn skriker, roper og slår, og slike hendelser utløses (eller betinges) vanligvis av andre familiemedlemmers atferd.

Allerede i 3-4 års alderen (eller tidligere) er det mulig å identifisere antisosial atferd eller forløperne til denne. Antall enkeltproblemer er den beste indikatoren på framtidig problematferd, både i førskolealder og i tidlig skolealder (Robins, 1966, 1978). Antisosial atferd er stabil over tid, og i en studie av 13 og 14 åringer fant Bendixen og Olweus (1999) en høy grad av rangeringslikhet med en korrelasjon på 0.80 over ett år og 0.60 over en toårsperiode. Antisosial atferd var mer stabil blant gutter enn blant jenter, noe som blant annet kan forklares med at jenter i mindre grad var innblandet i slike handlinger.

Undergrupper av alvorlige atferdsproblemer

Selv om barn og unge med alvorlige atferdsproblemer har mange likhetstrekk, er det også mulig ytterligere å skille mellom undergrupper som har visse felles kjennetegn eller felles utviklingsbaner (Loeber, 1990). Ett skille går mellom åpen aggressiv og utagerende atferd (eng. overt) og skjulte atferdsproblemer som å stjele, bedra, begå innbrudd og hærverk (eng. covert) (Loeber & Hay, 1994). Det er imidlertid stor grad av samvariasjon mellom disse formene for atferd, noe som har fått enkelte til å hevde at det dreier seg om et gradsspørsmål der de tyngst belastede har både åpne og skjulte problemer (Patterson, Dishion &

Chamberlain, 1993). Et annet skille går mellom barn som starter sin problemkarriere tidlig (eng. 'early onset' eller 'life course persistent') og de som først identifiseres i ungdomsalderen (eng. 'late onset' eller 'adolescent limited') (Moffitt, 1993). De som starter tidlig (dvs. før de er 10, eventuelt 12 år) har ofte en dårligere prognose for videre atferdsutvikling enn de som starter i ungdomsalder. Barn som debuterer tidlig og som senere blir vedvarende antisosiale viser ofte tegn på nevropsykologisk funksjonsproblemer (Moffitt, Caspi, Rutter & Silva, 2001). Funksjonssvikten antas å bidra til utvikling av vanskelig temperament, impulsivitet, hyperaktivitet, oppmerksomhetssvikt og lærevansker som igjen kan lede til en konfliktfylt tilpasning og utvikling av atferdsforstyrrelse. Forholdet som har blitt grundigst studert er svikt i utøvende (eng. executive) og verbal fungering, slik det kommer til uttrykk i AD/HD og i ulike lærevansker (særlig lesevansker). Når alvorlige atferdsproblemer først viser seg i ungdomsalder, er det bedre muligheter for rehabilitering, blant annet fordi disse ungdommene gjennom en mindre problembelastet barndom har hatt positive erfaringer med venner og familie, og har utviklet adekvate sosiale ferdigheter og mestringsferdigheter i forhold til stress.

Kjønnsforskjeller

Det er betydelige kjønnsforskjeller i forekomst og av alvorlige atferdsproblemer hos barn og unge og i et livsløsperspektiv er det ti ganger mer sannsynlig at gutter utvikler atferdsproblemer sammenlignet med jenter (Moffitt & Caspi, 2001). Men selv om det er betydelige nivåforskjeller mellom gutter og jenter, ser det ut til at jenter med alvorlige atferdsproblemer har en like dårlig langtidsprognose som gutter, og utviklingsforløpet har mange likhetpunkter. Bongers, Koot, van der Ende og Verhulst (2004) fant at foreldre rapporterte nesten dobbelt så mye aggressive atferd hos gutter som hos jenter i førskolealder, men at disse forskjellene nesten ble utlignet i ungdomsalder. Likhetspunktene i gutters og jenters antisosiale atferd ser ut til å være flere i ungdomsalder enn i noen annen utviklingsperiode. Epidemiologiske studier tyder på at jenter/kvinner som diagnostiseres med alvorlige atferdsvansker, som regel starter sin problemkarriere i ungdomsalder, og at deres problemer er mindre alvorlige og mindre kroniske (Frick, 1998). I ungdomsalder synes komorbiditet å være regelen heller enn unntaket for begge kjønn og de vanligste tilleggsprobemene er depresjon, angst, ADHD, lesevansker og rusproblemer (Moffitt, Caspi, Rutter & Silva, 2001). Kjønnsforskjeller i antisosial atferd varierer avhengig av om en studerer normative utvalg (Moffitt mfl., 2001) eller kliniske utvalg (Loeber, 1990; Loeber &

Hay, 1994), og forskningslitteraturen er ikke samstemt i synet på hva som er optimale behandlingstilbud for henholdsvis gutter og jenter (se Ogden & Amlund-Hagen, 2009).

Sosiale og kognitive ferdigheter

Alvorlige atferdsproblemer har både en problematferdsdimensjon og en sosial ferdighetsdimensjon, og selv om disse er negativt korrelerte, så representerer de likevel delvis uavhengige funksjonsområder med ulike determinanter (Sørli, Amlund-Hagen & Ogden, 2008). Selv om dimensjonene samvarierer, så har ikke alle antisosiale barn sosiale kompetanseproblemer. En betydelig andel av de aggressive barna kjennetegnes både av positiv og negativ sosial atferd, og de kan være både likt og mislikt (Price & Dodge, 1989). Antisosiale barn har imidlertid vanskelig for å lære seg atferdskontroll og grunnleggende sosiale ferdigheter som de trenger i sin videre tilpasningsprosess (Patterson, 1982). Dårlig sosial fungering henger sammen med aggresjon, vold, atferdsforstyrrelser og mobbing, men sosiale ferdighetsproblemer blir ofte oppfattet som sekundære følgetilstander av slike problemer (Beauchamp & Anderson, 2010). Selv om forskning så vel som klinisk praksis viser at utviklingen av sosiale ferdigheter er viktig, så har de en tendens til å bli oversett og nedprioritert i forhold til andre vansker, både av helsearbeidere og foreldre (men ikke av barna selv) (Beauchamp & Anderson, 2010). Dette begrunner hvorfor sosiale ferdigheter og sosial kompetanse bør inngå i utredning av alvorlige atferdsproblemer, men også i evalueringer av behandlingstiltak for denne målgruppen.

Kognitivt kjennetegnes ungdom med alvorlige atferdsproblemer av manglende eller sviktende informasjonsbearbeiding noe som kommer til uttrykk som lite funksjonelle sosiale problemløsningsferdigheter og tendenser til attribusjonsskjevheter (Crick & Dodge, 1994). Det vil si at de lett tillegger andre fiendtlige intensjoner, eller legger skylden for egne problemer på andre. Tvetydige eller uklare sosiale signaler blir lett tolket som aggressive og kan utløse negative reaksjoner. Sammenhengen mellom alvorlige atferdsproblemer og sosiale og kognitive ferdigheter underbygger at kartleggingen av slike ferdigheter også bør inngå i utredninger og evalueringer av tiltak (for eksempel Gresham & Elliott, 1990; Spence, 1995).

Tjenester og reaksjonssystem overfor alvorlige atferdsvansker

Samfunnets tiltak mot antisosialitet og kriminalitet har skiftet mellom straff, behandling, situasjonsforebygging, utdanning og rehabilitering. Behandlingsrettede eller psykososiale

tiltak kan fremme integrering og rehabilitering av sosialt utsatte barn og unge og slike tiltak har både sine tilhengere og kritikere (Lösel, 1993). Fra konservativt hold kritiseres behandling for å mangle en allmennpreventiv effekt mens de progressive kritikerne hevder at antisosialitet er uttrykk for et sosialt og samfunnsmessig problem - ikke en sykdom som kan behandles (Lösel, 1993). Tiltaks- og reaksjonssystemet overfor alvorlige atferdsvansker er i Norge preget av at vi har en relativ høy kriminell lavalder noe som fører til at strafferettslige reaksjoner er ganske uvanlig for barn under 15 år. Strafferettssystemet brukes også med varsomhet overfor ungdom i alderen 15-17 år (NOU 2008:15). Barn og unge som begår lovbrudd mens de er under den kriminelle lavalder kan bli varetektsfengslet, men saken blir deretter ofte oversendt til Barnevernstjenesten. Fengsling har i liten grad vist seg å forebygge ny kriminalitet. En gjennomgang av 150 straffesaker mot barn under 18 år som ble dømt til ubetinget fengsel i perioden fra januar 2006 til slutten av juni 2007 viste at mer enn 66% av ungdommene i løpet av kort tid hadde tilbakefall til ny alvorlig kriminalitet som medførte fengselsstraff (NOU 2008:15). Fengselsstraff er med andre ord ikke særlig egnet til å forebygge ny kriminalitet blant ungdom. Dette bekreftes også av en svensk etterundersøkelse av 470 ungdommer som ble dømt til fengsel (1991-1998) og 371 ungdommer som ble dømt til ”sluten ungdomsvård” (et lukket institusjonstiltak i regi av Statens Institusjonsstyrelse) (1999-2003). Oppfølgingsstudien som var basert på registerdata viste at av de som var dømt til fengsel hadde 60 prosent en ny dom og blant de som var dømt til ”sluten ungdomsvård” hadde 68 prosent en ny dom innen tre år hadde gått (Pettersson, 2010). En meta-analyse av internasjonale undersøkelser pekte i samme retning: strafferettslige sanksjoner/reaksjoner kunne ha en moderat forebyggende effekt, men førte i verste fall til økt tilbakefall mens rehabiliterende behandling viste seg å redusere tilbakefallsprosenten hos dømte lovbytere (Lipsey & Cullen, 20007).

Storvoll gjennomførte i 1997 en kartleggingsundersøkelse som viste et mangfoldig reaksjons- og tjenestesystem overfor barn og unge med alvorlige atferdsproblemer. Både PP-tjenesten, Barneverntjenesten og Psykisk helsevern (Barne- og ungdomspsykiatrien) kartlegger/utredet barn og unge med alvorlige atferdsproblemer. De ulike etatene samarbeider mye om disse sakene, men tiltakene var som oftest forankret i den enkelte etat. PP-tjenesten i grunnskolen gir råd og veiledning til familien mens Barnevernstjenesten både kan gi råd, iverksette tilsyn med hjemmet, gi tilbud om støttekontakt, avlastning, barnehageplass eller økonomisk tilskudd til familien/barnet (hjelpetiltak). Barnevernet kan også plassere barn/ungdom på institusjon eller i fosterhjem (omsorgstiltak). Barne- og ungdomspsykiatriske poliklinikker kan gi tilbud som omfatter korttids individualterapi, familierapi og

medisinering. I skolen ble det gitt spesialundervisning og oppfølging med bruk av assistenter. De fleste ungdommer som fikk hjelp i rusmiddelomsorgen fikk døgnerbehandling (67%), og blant institusjonsalternativene var kollektiv det mest brukte. Flere av instansene ga uttrykk for at det var vanskelig å gjøre noe når barna/ungdommene selv ikke var motiverte eller ønsket hjelp, og når støtten fra familien var liten. Institusjonsplasseringer ble betraktet som det mest aktuelle tiltaksalternativet og flere påpekte at det var mangel på slike (Storvoll, 1997).

Storvolls (1997) kartleggingen av tiltak overfor barn og unge med alvorlige atferdsproblemer avdekket flere svakheter ved tiltakssystemet. For det første var det vanskelig å skaffe oversikt over situasjonen, for eksempel hvor utbredt problemene var og hva behovene var. Registreringsrutiner og offentlig statistikk var mangelfulle og ufullstendige, og det er vanskelig å dokumentere graden av overlapping i klientgruppene mellom psykiatri, barnevern og pp-tjenesten. For det andre viste kartleggingen en betydelig mangel på virksomme tilbud, og at tilbudene ikke sto i forhold til problemenes alvorlighetsgrad. For det tredje var det vanskelig å vite hva som styrte prioriteringer og etterspørsel etter tilbud. Beskrivelsen av prioriterte behov varierte med hvem i tjenesteapparatet som foretok vurderingene og de kunne skifte over tid. De ulike barne- og ungdomstjenestene ga ulike signaler og disse ble igjen påvirket av faglige stemningsskifter i synet på problemenes innhold og utbredelse og på hvilke tiltak som var adekvate. I tillegg foregikk samarbeidet i et tverrfaglig og tverretattlig spenningsfelt som skapte en rekke avgrensings- og samarbeidsproblemer i selve tiltaksapparatet. Potensialet for konflikter mellom instansene som arbeidet med de vanskeligste samt mellom etater og forvaltningsnivåer, var betydelig (Storvoll, 1997). Selv om kartleggingen ble foretatt for mer enn ti år siden, er det lite som tyder på at problemene som Storvoll (1997) beskrev er løst. Det framgår blant annet av en utredning som 12 år senere drøftet samarbeidet mellom kommunale tjenester og mellom kommunale tjenester og statlige tjenester som gir hjelp til barn og unge og deres foreldre .

Det såkalte Flatø utvalget (NOU 2009: 22) peker på samarbeidsproblemer som blant annet skyldes at ingen har et formelt koordineringsansvar overfor brukerne. Tjenestene beskrives i stor grad som fragmenterte (ulike etater og forvaltningsnivåer) slik at det lett oppstår lett brudd i overgangssituasjonene som for eksempel mellom barnehage og skole. Kvaliteten på hjelpen er ofte mer et resultat av enkeltpersoners innsats enn av et velfungerende system. Utvalget peker på at terskelen for tilgang til psykisk helsehjelp oppleves som høy og det mangler systemer for tidlig intervensjon. Det synes videre å være et spenningsnivå mellom forvaltningsnivåene, blant annet mellom kommunalt og statlig barnevern. Samarbeidet mellom barnvernstjenesten og statlig psykisk helsevern er også

problematisk på grunn av ulikt utdanningsnivå og til dels ulike teoretiske forankringer. Regelverket blir ofte tolket unødig innskrenkende og synes ikke å gi en felles plattform for flerfaglig samarbeid. Det fører til at de ulike deltjenestene i stor utstrekning bare tar ansvar innenfor eget område og at samarbeidet begrenser seg til det som skal til for å fremme egen oppgaveløsning. Finansieringssystemet synes også å hemme samarbeid (blant annet refusjonstakstene i psykisk helsevern og kommunal medfinansiering på andre områder). Det er også mulig at ressursknapphet fører til at tjenestene blir opptatt av å avgrense egne oppgaver og skyve fra seg problemer som andre instanser kan ta (NOU 2009:22).

Uavhengige av hvilke tiltak som iverksettes, så har det vist seg at alvorlige atferdsproblemer er vanskelige å endre; behandlingen er utfordrende og hindringene mange. Det viser blant annet Kjellsberg og Dahls (1998) etterundersøkelse av 1095 barn og unge som hadde vært innlagt ved Statens senter for barne- og ungdomspsykiatri i Norge i perioden 1963-81. I gjennomsnitt hadde det gått 25 år siden innleggelsene og disse hadde i gjennomsnitt vært 5.5 måneder. Den vanligste diagnosen var alvorlige atferdsforstyrrelser (eng. disruptive behaviour disorder) som 45 prosent hadde og av disse hadde 36,5 prosent også rusproblemer (eng. psychoactive substance use disorder). De tidligere klientene ble fulgt opp gjennom offentlige registerdata og på oppfølgingstidspunktet fant forfatterne igjen 52 prosent av klientene i strafferegisteret mens 38,5 prosent var i trygderegisteret og 13,6 prosent var døde. Det var bare 22,9 prosent av som hadde klart seg bra i den forstand at de ikke befant seg i noen av registrene, og av disse var 2/3 kvinner. Formulert på en annen måte hadde 33,3 prosent av kvinnene og 13,9 prosent av mennene klart seg bra. Selv om de langsiktige oppfølgingsresultatene var lite oppmuntrende, er det vanskelig å vite om resultatene hadde vært enda dårligere uten behandlingen fordi undersøkelsen manglet en kontrollgruppe (Kjellsberg & Dahl, 1998). Resultatene bidro likevel til å at det ble stilt spørsmål ved verdien av institusjonsbehandling av alvorlige atferdsproblemer. Et viktig bidrag til denne debatten kom også fra Sverige, der Vinnerljung, Sallnäs og Westermarck (2001) dokumenterte at foruroligende mange tenåringsplasseringer i fosterhjem og på institusjon brøt sammen, og at dette særlig gjorde seg gjeldende for ungdom med antisosial atferd. I en undersøkelse av alle plasseringer av 13-16 åringer i Sverige fra 1991 (N = 1110) viste det seg 37 prosent av alle plasseringer av barn og unge med antisosial atferd brøt sammen innen 5 år sammenlignet med 25 prosent av de øvrige plasseringer (Vinnerljung mfl., 2001).

Et eksempel på langsiktige negative virkninger av psykiske lidelser generelt og alvorlige atferdsproblemer spesielt framgår også av en rapport om veksten i antall

uføretrygdede i Norge (Econ rapport, 2009). I perioden 2000-2009 økte antall unge uføretrygdede i alderen 18-29 år med 41 prosent (mot 24% i alle aldersgrupper). Seks av ti unge uføre hadde en psykiatrisk diagnose, inkludert atferds- og personlighetsforstyrrelser (bl.a. ADHD) og blant disse var det nesten en tidobling av andelen uføretrygdede. Rapporten konkluderte med at endringer i arbeidslivet hadde gjort det vanskeligere for unge med psykiske sykdommer å skaffe seg og beholde en jobb, og det var særlig økte krav til kommunikative ferdigheter og sosial kompetanse i alle typer jobber som var kritiske. Økte krav til sosiale ferdigheter i arbeidslivet samt flere, hyppigere og tidligere diagnoser ble framhevet som de viktigste forklaringene på veksten i antall unge uføre. For de med atferds- og personlighetsforstyrrelser kan det dessuten ha spilt negativt inn at svært mange ikke fullførte videregående skole, noe som begrenset deres muligheter både i utdanning og arbeidsliv (Econ rapport, 2009).

I sin beskrivelse av endringene i tjenester og tiltak det siste tiåret framhever Flatø utvalget (NOU 2009: 22) at mange spesialskoler og institusjoner har blitt nedlagt mens en har forsøkt å styrke tilbudet til barn og unge lokalt i barnets hjemmemiljø. Det legges stor vekt på betydningen av tidlig innsats og selv om tjenester til barn og unge fortsatt har et individperspektiv, er det i tillegg mer fokus på systemrettet arbeid. Det pekes også på at det er innført nye forskningsbaserte metoder i arbeidet med barn og unge (for eksempel De utrolige årene, PMTO og MST). Når det gjelder framtidig satsing peker utvalget på viktigheten av klare og realistiske mål, klart definerte roller, sterk ledelse, styring på tvers av tjenester, og gode systemer for deling av informasjon (NOU 2009: 22).

Det kan også spores en viss optimisme i forskningslitteraturen når det gjelder mulighetene for å redusere alvorlig problematferd gjennom en kombinasjon av tiltak i hjem, skole og nærmiljø. Loeber og Farrington (1998) formulering om at «det aldri er for tidlig å forebygge og aldri for sent å behandle», synes fremdeles å ha gyldighet.

Oppsummerende diskusjon

Alvorlige atferdsproblemer kan defineres utviklingspsykologisk som grader av antisosial atferd eller medisinsk-diagnostisk som atferdsforstyrrelser. Det dreier seg om et flerdimensjonalt problem som varierer i forhold til kjønn, alder, tidspunkt for når problemene viser seg og i forhold til forløp og alvorlighetsgrad. Alvorlige atferdsproblemer er et sosialt definert problem og siden inklusjons- og avgrensningskriteriene endrer seg over tid er det

vanskelig å finne klare prevalenstall. Men et rimelig estimat synes å være at cirka 3-5% av et årskull har problemer som kvalifiserer for behandling eller andre tiltak.

Synet på hva hvilke tiltak som bør iverksettes overfor barn og ungdom med alvorlige atferdsproblemer har skiftet over tid, og mens institusjonsplassering og fengsling i økende grad har kommet i miskreditt, har åpen psykososial og rehabiliterende behandling i familie- og nærmiljø fått økt legitimitet gjennom lovende forskningsresultater. Historisk sett kan utviklingen beskrives som en bevegelses fra naiv optimisme til mørk pessimisme og deretter til nøktern realisme i synet på å redusere alvorlige atferdsproblemer gjennom behandling og andre psykososiale tiltak. I de senere årene har forskningen fått en mer sentral plass som premissleverandør for utviklingen av praksis, men intervensjons- og evalueringsforskningen har i større grad fokusert på utviklingen av virksomme metoder og tiltak enn på de strukturelle forholdene som skaper problemer for samarbeidet på tvers av sektorer og etater. Et tiltaksapparat som er sektor- og etatsinndelt med ulikt lovgrunnlag og ulike finansieringsordninger, skaper barrierer for samarbeid og samhandling mellom tjenester og organisasjoner. Det knytter seg problemer til et fragmentert tiltaksapparat som har samarbeidsproblemer, men også prioriterings og motivasjonsproblemer. Det er ikke fristende å arbeide med barn og unge som åpent gir uttrykk for at de ikke ønsker hjelp, og som aktivt motarbeider velmente forsøk på å få dem til å fungere sosialt i skole, hjem og fritid. Ofte kommer hjelpen sent, når problemene har vokst familie og omgivelser over hodet, og altfor ofte har innsatsen vært lite treffsikker og har gitt dårligere resultater enn forventet.

Selv om det nå synes akseptert at åpen behandling eller åpen behandling i kombinasjon med plasseringstiltak er best egnet til å forebygge videre problemutvikling og tilbakefall til kriminalitet, så har plasseringstiltakene i form av fengsling og internering fortsatt å spille en sentral rolle blant tiltakene for barn, men særlig for ungdom med alvorlige atferdsproblemer. Før 1990 var programbegrepet relativt ukjent i barne- og ungdomstjenestene i Norge der arbeidet som regel hadde referanser til behandlingsmodeller (for eksempel miljøterapi) eller arbeidets teoretisk orientering. Det psykodynamiske perspektivet ble ofte satt opp mot det læringsteoretiske perspektivet, selv om det også fantes eksempler på at disse kunne integreres i modeller for miljøterapeutisk institusjonsbehandling (Vedeler, 1973; Selnes & Larsen, 1974). Men ved i løpet av 1990 årene hadde evidensbaserte programmer og praksis blitt et etablert begrep i Norge gjennom internasjonal påvirkning og samarbeid, blant annet med fag- og forskningsmiljøer i USA. I 1997 syntes tiden å være moden for en ny policy og nye initiativ der det ble satset sterkere på familie- og nærmiljøbaserte tiltak som PMTO og MST.

2. Evidensbasert policy og utviklingen av evidensbaserte behandlingsprogrammer

Introduksjonen av evidensbaserte behandlingsprogrammer i Norge skjedde i en periode med økende interesse for og diskusjon rundt temaet evidensbasert politikk og praksis. Temaet ble drøftet innenfor områder som klinisk medisin og psykologi, pedagogikk, sosialt arbeid og kriminologi. Systematiske forsknings- og kunnskapsoversikter ble utarbeidet av Cochrane Collaboration som ble opprettet i 1993, og etter hvert av andre som Campbell Collaboration og EPPI senteret (the Evidence for Policy and Practice Information and Coordinating Centre). I Norge ble denne utviklingen møtt både med interesse og skepsis, ikke minst fordi randomiserte kontrollerte studier ble framhevet som 'gullstandarden' for evalueringer av virksomme tiltak. Som en reaksjon på de strenge kravene til forskningsdesign, kom begrepet 'evidence-based policy' til å stå for et bredere og mer inkluderende perspektiv på 'evidens' enn det som kom til uttrykk gjennom de presist definerte kravene til 'empirically supported programs'.

'Evidensbasert policy' kan defineres som en tilnærming som : "... helps people make well-informed decisions about policies, programmes and projects by putting the best available evidence from research at the heart of policy development and implementation" (Davies, 2004, s. 3). Nutley, Morton, Jung and Boaz (2010) beskriver utviklingen av den evidensbaserte tilnærmingen til politikk og praksis i Storbritannia som et forsøk på å oppmuntre politikere, beslutningstakere og praktikere til å anvende den beste tilgjengelige kunnskap i form av robust og forskningsbasert evidens og å verdsette kvalitet og tydelige resultater. Samtidig erkjenner de at det er ulike oppfatninger av hvilken status og betydning ulike former for evidens har og hvordan evidens skal anvendes i utformingen av politikk og praksis (Nutley mfl., 2010).

En annen forståelse hadde Fixsen, Naoom, Blasé, Friedman & Wallace (2005) som beskrev den evidensbaserte bevegelsen (eng. evidence based movement) som et internasjonalt initiativ for å knytte en sterkere forbindelse mellom forskning, policy og praksis (Fixsen mfl., 2005). Deres forståelse knyttes nærmere opp til formidling og implementering av

intervensjonsprogrammer som bygget på og hadde vist seg å være virksomme gjennom forskningsbaserte evalueringer. Evidensbasert policy har i tråd med denne forståelsen også blitt beskrevet som en økt interesse for, bevissthet om og bruk av evidensstøttede behandlingsformer i praksis (Chambless & Ollendick, 2001). Det som skiller den britiske (Nutley mfl., 2010) fra den nord amerikanske (Fixsen mfl., 2005) forståelsen av evidensbasert policy er spørsmålet om hvor forpliktende forholdet mellom forskning og praksis skal være. Og mens utviklingen i Storbritannia har vært nært knyttet til politiske beslutninger har utviklingen i USA i stor grad handlet om utviklingen av evidensbaserte intervensjoner.

'Evidensbasert policy' slik den beskrives i Storbritannia (Davies, 2004; Nutley mfl., 2007, 2010) handler om å benytte forskning i vid forstand om å ta beslutninger samt utvikle og iverksette politiske eller administrative planer i praksis. Det dreier seg om å formulere helhetlige og langsiktige mål og la forskning legge premisser for tiltak og tjenester. Men policy kan også dreie seg om beslutninger som påvirker prevalensen av et problem uten bruk av individrettede intervensjon som når en forsøker på å redusere forbruket av alkohol og tobakk gjennom å øke avgifter eller redusere tilgjengelighet (Biglan & Ogden, 2008).

Mens evidensbasert praksis har fanget praktikernes oppmerksomhet har evidensbasert policy først og fremst blitt et tema for dem som beskjeftiger seg med politikk og forvaltning. Policy kan utformes på nasjonalt, regionalt, kommunalt eller lokalt plan av politikere, av forvaltningen eller av organisasjoner. Det kan dreie seg om et departement, et direktorat, kommunale eller regionale organisasjoner (eks. helseforetakene). Disse aktørene er forståelig nok opptatt av hvordan de kan iverksette reformer og utviklingsarbeid som har en stor grad av faglig troverdighet og som med stor grad av sannsynlighet gir gode resultater og bidrar til at de når sine overordnede mål til en rimelig kostnad.

Det kan skilles mellom et "bredt" og et "smalt" perspektiv på forholdet mellom forskning, policy og praksis der det brede perspektivet definerer en løsere forbindelse mellom nivåene mens det smale perspektivet beskriver et mer forpliktende samspill mellom de tre.

Et bredt og et smalt perspektiv på sammenhengen mellom policy, forskning og praksis

Det 'brede' perspektivet har særlig preget synet på forholdet mellom policy, forskning og praksis i Storbritannia og forklarer hvordan forskning kan påvirke praksis på langsiktige, indirekte og nyanserte måter (Nutley mfl., 2010). Dette kalles også *konseptuell* bruk av forskning og bidrar til å gjøre planleggere og praktikere mer bevisste med hensyn til

beslutningene de tar. Konseptuell (i motsetning til instrumentell) bruk av forskning vil si at den informerer og inspirerer policy og praksis, eller bidrar til diskusjon og debatt (Nutley mfl., 2007). Dette omtales også som *opplysningsmodellen* (eng. enlightenment; Weiss, 1979) fordi forskning påvirker hvordan man tenker om problemer og deres løsninger, og hvordan den påvirker beslutninger og policy, gradvis og kumulativt. Selv om dette kanskje er det vanligste perspektivet på forholdet mellom forskning og praksis også i Norge (Ogden, Kärki & Teigen, 2010) så kan det innvendes mot opplysningsmodellen at den mangler prosedyrer og kriterier for å filtrere ut forskning som er foreldet eller som er av dårlig kvalitet (Weiss, 1979). Forskningen kan derfor like gjerne føre til 'mørklegging' som til 'opplysning'. Med andre ord kan opplysningsmodellen føre til *strategisk bruk* av forskning hvis den brukes til å underbygge beslutninger som allerede er tatt på et annet grunnlag (Nutley mfl., 2007). Brukt på denne måten kan forskning brukes til å underbygge etablerte standpunkter eller å utfordre andres. Kunnskapen er ikke nødvendigvis usaklig, men blir tilpasset visse interesser eller ideologier; og så fort én part belegger sitt syn med forskningskunnskap, kan motparten etterspørre alternativ forskningskunnskap for sitt formål (Hatland & Bay, 2009). I et komplisert saksfelt hvor det er stor politisk uenighet om målene, vil en kunne forvente strategisk bruk av forskning for slike taktiske formål.

Den brede definisjonen av evidensbasert policy forutsetter at politiske beslutninger bygger på forskningsbasert evidens, men også på profesjonelle normer, verdier, ideologi og politisk overbevisning. Forskningskunnskap må derfor konkurrere med personlige erfaringer og vurderinger hos beslutningstakerne, ekspertsynspunkter, konsulentrapporter, vaner og tradisjoner i praksisfeltet samt innspill og påvirkninger fra interesse- og pressgrupper eller media (Davies, 2004).

Etablering av evidensbasert policy kan også hemmes av at kunnskapsstatus på visse fag- og problemområder er så mangelfull at det ikke er grunnlag for å foreta evidensbaserte beslutninger. Et annet kompliserende forhold kan være at budskapet fra forskningen av og til er motsetningsfylt, uklart og preget av mange forbehold. Videre kan forskning ha uklare mål, svake forskningsdesign, metodologiske svakheter, ufullstendige statistiske analyser, selektiv bruk og rapporteringer av data, og konklusjonene støttes ikke alltid av data. Siden enkeltstudier sjelden kan gi et representativt bilde av kunnskapsstatusen på sitt område legger det brede policyperspektivet mer vekt på systematiske kunnskapsoversikter eller meta-analyser som er mindre utvalgs-, tids- og kontekstspesifikke og som formidler hvilke muligheter og begrensninger generaliseringene har (Davies, 2004).

Det 'smale' perspektivet på evidensbasert policy omtales også som den evidensbaserte tilnærmingen der intervensjoner, programmer eller tiltak bygger på og valideres gjennom forskning. Forskning brukes i denne sammenhengen direkte for å endre praksis for en bestemt målgruppe og i arbeidet med et definert problem. Resultater fra intervensjonsstudier kan omsettes til forpliktende intervensjonsprogrammer, verktøy og retningslinjer for praksis. Både eksperimentelle primærstudier og systematiske kunnskapsoversikter eller meta-analyser kan være drivkraften bak evidensbaserte modeller, program og metoder. Et slikt perspektiv forutsetter imidlertid problemløsende forskning som kan være relativt uavhengig av interesser og verdier (Hatland & Bay, 2009). Innenfor dette perspektivet er det ofte *konsensus* mellom forskere og beslutningstakere om hva som er viktige problemer og hvordan disse bør løses. Videre forutsetter en at det er enighet om metodene for kunnskapsinnhenting og også om hvordan kunnskapen skal brukes for å forbedre tjenestetilbudet og dermed utbyttet for brukerne. Noen omtaler dette som instrumentell bruk av forskning, og der målet er mer effektive beslutninger og et mer resultateffektivt tjenestetilbud, blant annet i form av evidensbaserte programmer (Nutley mfl., 2007). Forskeren er i denne sammenhengen både premissleverandør og aktiv bidragsyter i utviklingen av praksis, og kan fungere i rollene som både programutvikler og evaluatør. Heller enn å forske *på* praksis dreier det seg om å forske for å *utvikle* praksis. Dette står i motsetning til en mer kritisk og distansert tilnærming der forskeren heller enn å engasjere seg og delta, inntar rollen som observatør av beslutninger og vurderer alle tiltak i et kritisk perspektiv. Den kritiske forskerrollen tar utgangspunkt i at alle policy beslutninger, planer og prosjekter har sine fordeler og ulemper og at forskningens oppgave er å utfordre de teoretiske, empiriske eller analytiske svakheter i disse, enten de er tilsiktet eller ikke (Davies, 2004; Nutley mfl., 2007).

Storbritannia har den lengste tradisjonen for det brede perspektivet på evidensbasert policy mens det smale perspektivet særlig har preget feltet som omtales som tidlig intervensjon i forhold til risikoutsatte barn, blant annet "Head Start" programmet i USA. Arbeidet med å utvikle evidensbaserte forebyggings- og behandlingstiltak overfor alvorlige atferdsproblemer hos barn og unge i Norge kan også sies å ha utspring i en nasjonal policy i tråd med konsensusperspektivet (Ogden, Amlund-Hagen, Askeland & Christensen, 2009). Et annet eksempel på nasjonal policy utforming var den svenske regjeringens oppdrag til Socialstyrelsen i 1999 om å utvikle en strategi for forskningsbasert kunnskapsutvikling innen sosialtjenesten. Gjennom flere år har derfor svenske praksis- og forskningsmiljøer, særlig Institutt for Metoder i Sosialt arbeid (IMS) arbeidet med å utvikle en kunnskapsbasert praksis (Jergeby, 2008; Olsson, 2008; Tengvald, 2001). Målsettingen har vært å skape en struktur for

systematisk kunnskapsutvikling og effektiv formidling av informasjon slik at offentlige sosiale tjenester i større grad har blitt evidensbasert (Socialstyrelsen, 2000).

Utviklingen av evidensbaserte behandlingsprogrammer

Prinsippet om at behandling eller tiltak skal bygge på forskning og evalueres gjennom kontrollerte evalueringsstudier er ikke nytt i internasjonal sammenheng. På 1960 tallet begynte forskere innen psykologi og pedagogikk å rapportere om arbeidsmåter som gjennom kontrollerte evalueringer viste seg å gi bedre resultater enn metodene de ble sammenlignet med. For eksempel viste Wesley Becker og kolleger (Becker, Madsen, Arnold & Thomas, 1967; Madsen, Becker & Thomas, 1968) i et tidsseriedesign at læreres bruk av regler, ros og ignorering kunne bidra til signifikante forbedringer av problematferd i skoleklasser. I årene som fulgte rapporterte forskere om intervensjoner som kunne løse flere viktige psykologiske og atferdsmessige problemer fra enurese til sigarett røyking, fra depresjon til ekteskapskonflikter (Kazdin, 1978). Det er imidlertid først de siste par tiårene at slik forskning har blitt lagt til grunn for det som kalles evidensbasert forebygging, behandling eller praksis.

Interessen for evidensbasert praksis fikk i løpet av 1990 årene et oppsving gjennom introduksjonen av *evidensbasert medisin* (Sackett, Rosenberg, Gray, Haynes & Richardson, 1996; Sackett, Richardson, Rosenberg & Haynes, 1997) og dette påvirket igjen utformingen av prinsippene for evidensbasert psykologisk praksis i den amerikanske psykologforeningen (APA, 2006). Den norske psykologforeningen har også gitt sin tilslutning til disse prinsippene (Rønnestad, 2008). Sackett og kolleger (1997) beskrev evidensbasert medisin som en integrering av tre kunnskapskilder: (1) den for tiden beste vitenskaplige kunnskapen, (2) brukernes erfaringer og ønsker og (3) den profesjonelles erfaringer og kompetanse. Vitenskaplig kunnskap dreide seg om kontrollerte evalueringsstudier (effektstudier) der en sammenlignet grupper som fikk et behandlingstiltak (intervensjons- eller eksperimentgruppen) med grupper som fikk andre tilbud (kontrollgrupper). Praktikere skulle bidra gjennom å samle inn informasjon fra klienten selv eller andre, vurdere informasjonen og integrere forskningskunnskap om mulige tiltak med klientens ønsker og egne praksiserfaringer. Når tiltaket ble gjennomført, var det praktikerens oppgave å evaluere resultatet for klienten eller pasienten. Prinsippene for evidensbasert medisin ble også lagt til grunn for evidensbasert praksis i svensk sosialtjeneste og definert som: "... en noggrann, öppet redovisad och omdömesgill användning av den för tillfället bästa evidensen för

beslutsfattande om insatser till enskilda individer” (Jergerby & Sundell, 2008, s. 12). Dersom det manglet vitenskaplig grunnlag for å intervenere, eller det finnes flere ekvivalente intervensjoner, ville brukernes erfaringer og ønsker samt den profesjonelles erfaringer og kunnskaper få stor betydning for valg av tiltak. Evidensbasert praksis skal videre være transparent slik at det er mulig å synliggjøre motivene for beslutninger, og å kunne innta en kritisk holdning til kunnskap (Jergerby & Sundell, 2008). Mens utviklingen av evidensbasert praksis ble påvirket av evidensbasert medisin slik den ble definert av Sackett mfl. (1996) og av evidensbasert psykologisk praksis (Chambless & Ollendick, 2001) så kom utviklingen av evidensbasert behandling til å følge en annen linje der en stilte mer eksplisitte krav til behandlingens innhold og kriteriene for evidens. Både den amerikanske psykologforeningen (APA, 2006), det såkalte Blueprint initiativet (Elliott, 1998) og organisasjoner som Society for prevention research (www.preventionscience.org) (Flay mfl., 2005) har definert hvilke forskningskrav som bør stilles til evidensbaserte program eller ’empirically supported therapies’. Dette er tema for neste avsnitt.

Kriterier for evidensbaserte programmer

Standardene for evidens er etablerte konvensjoner for hvilket innhold behandlingen skal ha og hvilke krav som skal stilles til forskning som behandlingen bygger på. En arbeidsgruppe i Divisjon 12 for klinisk psykologi i den amerikanske psykologforeningen (APA) gjennomgikk empirisk støttede tiltak (EST) og forebyggende program for barn (the Task Force on Effective Psychosocial Interventions: A life-span perspective) og skilte mellom behandlingstiltak basert på følgende kriterier (Chambless mfl., 1998):

- (I) *Veletablert behandling* (eng. well established treatments) Disse skal være testet i minst to eksperimenter med gruppedesign som viser effekt i form av (a) bedre resultater enn medikamentell eller psykoterapeutisk placebo eventuelt annen behandling eller (b) ekvivalens i forhold til allerede etablerte behandlinger med adekvate utvalgsstørrelser. Alternativt kan de være testet i en serie eksperimenter med enkeltkasus design (eng single-case) som viser effekt sammenlignet med en annen form for behandling. Utvalgets kjennetegn skal spesifiseres og eksperimenter skal gjennomføres med bruk av behandlingsmanualer, eller tilsvarende klare beskrivelser av behandlingen. Når en har dokumentert behandlingseffekt i minst to studier gjennomført av uavhengige forskerteam, kan behandlingen beskrives som ”*efficacious*”. Hvis det bare foreligger én studie, eller

hvis alle studiene er utført av det samme forskerteamet, så omtales behandlingen som "*possibly efficacious*" i påvente av replikasjoner.

- (II) *Antagelig effektiv behandling* (eng. probably efficacious treatments). Her stilles det krav om to eksperimenter som viser at behandlingen gir bedre resultater enn en kontrollgruppe på venteliste, eller to eksperimenter som tilfredsstiller kriteriene for bedre resultater enn medikamentell eller psykoterapeutisk placebo eventuelt annen behandling eller ekvivalens i forhold til allerede etablerte behandlinger med adekvate utvalgsstørrelser. Alternativt skal behandlingen være testet i en mindre serie med eksperimenter med enkeltkasus design og som tilfredsstiller etablerte behandlingskriterier.
- (III) *Eksperimentell behandling* (eng. experimental treatments) er behandlinger som foreløpig ikke er testet i studier som tilfredsstiller arbeidsgruppens metodekrav.

Flay mfl (2005) formulerte på oppdrag av Society for Prevention Research (SPR) standarder for programmer og policy for effektiv forebygging. Deres forslag til begrepsbruk har mange likhetpunkter med de som ble utarbeidet av arbeidsgruppen nedsatt av APA, og dekker to intervensjonsnivåer som tilsvarer studier av effekt og effektivitet. '*Efficacious*' intervensjoner har gjennom eksperimentelle evalueringer vist seg å ha signifikant effekt når forskerne har tett kontroll med implementeringen og mer ressurser enn det som er vanlig for å gjennomføre tiltaket. Det skal foreligge minst to undersøkelser som (1) omfatter definerte utvalg fra definerte populasjoner, (2) anvender psykometrisk holdbare mål og prosedyrer for datainnsamling, (3) analyserer data med rigorøse statistiske tilnærminger, (4) viser konsekvent positive effekter (uten betydelige negative og utilsiktede effekter) og (5) rapporterer minst én signifikant langtids oppfølging. Det andre nivået beskriver '*effective*' intervensjoner som har gitt nærmere spesifiserte resultater i eksperimentelle evalueringer i ordinær praksis. Effektive intervensjoner skal tilfredsstille kriteriene for effekt studier, men skal i tillegg ha (1) manualer, egnet opplæring, og teknisk støtte slik at en tredjepart kan adoptere og implementere intervensjonen, (2) evaluering under vanlige praksisbetingelser i undersøkelser med gode målinger av implementeringsnivå og målgruppens engasjement (både i intervensjons- og kontrollbetingelser), (3) angi hvilken praktisk betydning resultatene har, og (4) vise klart hvem resultatene kan generaliseres til (Flay mfl., 2005).

Kriteriene for evidens som ble utarbeidet av Flay mfl. (2005) retter i større grad enn APA-kriteriene fokus mot forskningskrav til testing av programmer i ordinær praksis. Det handler om økt oppmerksomhet i forhold til implementeringsgrad og implementeringskvalitet.

Kriteriene for effektive intervensjoner er derfor de mest aktuelle når programmer som har blitt evaluert i effektstudier skal spres til nye utprøvningssteder. Slike intervensjoner skal i tillegg til å være virksomme i forhold til målgruppens problemer være forhåndsbeskrevet og støttet av en implementeringsorganisasjon som har tilstrekkelig kapasitet og kompetanse. Videre skal programmene formidles på en slik måte at en både kan evaluere hvor godt de blir implementert og hvor virksomme de er på nye utprøvningssteder. Disse kriteriene sammenfaller også med kravene til at programmer som skulle inkluderes i den såkalte Blueprint serien.

Blueprint prosjektet er ett av de mest omfattende forsøkene på å identifisere modell program som har vist seg virksomme i arbeidet med å redusere voldelig kriminalitet, aggresjon, norm- og regelbrudd og rusmisbruk blant ungdom (Elliott, 1998). Av de mer enn 600 programmene som har blitt gjennomgått har 11 blitt beskrevet som ”effektive” mens 18 har blitt beskrevet som ”lovende”. Kravene til effektive eller virksomme program er at de skal ha et sterkt forskningsdesign (randomisering av deltakere eller kvasi-eksperimentelle design med matchede kontrollgrupper) med evidens for at de reduserer vold, kriminalitet eller rusmisbruk, resultater som vedvarer minst ett år etter at programmet er avsluttet og replikasjon med positive resultater på mer enn ett utprøvningssted (Elliott & Mihalic, 2004). Det ble også stilt krav om at programmene skulle ha tilstrekkelig formidlingskapasitet til at interesserte miljøer og organisasjoner kunne få hjelp og støtte fra programutviklerne til å implementere programmet på nye steder. Da serien ble etablert, ble Multisystemisk terapi (MST) inkludert som et av de effektive programmene mens Oregon modellen av PMT på det tidspunktet bare hadde vært benyttet i rene forskningsprosjekter. Det var med andre ord ikke etablert noen implementeringsorganisasjon, og samarbeidet med Atferdssenteret i Norge ble det første forsøket på å formidle behandlingsprogrammet i stor skala (Ogden, Forgatch, Askeland, Patterson & Bullock, 2005).

Etter hvert som det har blitt utviklet et betydelig antall evidensbaserte program har spørsmålet om hvordan en formidler slike program kommet mer i fokus. Implementeringsforskning (også omtalt som translational research) har særlig fokusert på prosessen der programmer som har vist seg virksomme i kontrollerte evalueringstudier spres eller formidles til nye utprøvningssteder og målgrupper (Elliott, 1998; Fixsen mfl., 2005).

Begrepsmodeller for implementering av evidensbaserte programmer

Implementering er et målrettet og systematisk arbeid for å omsette et virksomt program eller tiltak i ordinær praksis og implementeringsforskningen identifiserer hvilke forhold som er hemmer eller fremmer denne prosessen (Fixsen mfl., 2005; Sørli, Ogden, Solholm & Olseth, 2010). Utvikling, implementering og evaluering av evidensbaserte programmer er i fag- og forskningslitteraturen beskrevet gjennom begrepsmodeller som gir en oversikt over hvilke komponenter som er viktige for en vellykket implementering av evidensbaserte intervensjonsprogrammer (Biglan & Taylor, 2000; Burns, 2003; Fixsen mfl., 2005; Glasgow, Vogt & Boles, 1999; Kellam & Langevin, 2003; Price & Lorion, 1989). En variant av dette finnes i CONSORT statement (Moher, Schulz & Altman, 2001) som beskriver hvilke rapporteringskrav som bør stilles til randomiserte kontrollerte studier (jf kapittel 4). En litt annen innfallsvinkel har Kellam og Langevin (2003) som presenterte en oversikt over forskningsfaser i utviklingen av evidensbaserte programmer og de formulerte hvilke krav som bør stilles til et program etter at det har blitt testet i effektstudier og skal implementeres i ordinær praksis i liten (effektivitetsstudier) eller stor skala ('going-to-scale).

Forskningsfaser i utviklingen av evidensbaserte programmer

Kellam og Langevin (2003) formulerte fem faser i forskningen om evidensbaserte forebyggende intervensjoner: Fase 1: *efficacy trials* er effektstudier som tester om intervensjoner virker under optimale betingelser, fase 2: *effectiveness trials* er effektivitetsstudier som under ordinære praksisbetingelser tester for hvem, under hvilke betingelser og til hvilke kostnader intervensjoner virker, fase 3: *sustainability trials* er vedlikeholdsundersøkelser som tester hvilke opplæringstiltak som trengs for å vedlikeholde god praksis og hva kostnadene vil være, fase 4: *going-to-scale trials* er storskala implementeringsstudier der en tester modeller for å implementere programmer på bred front i lokalmiljøer og institusjoner/organisasjoner samt kostnadene ved disse modellene og fase 5: *sustainability systemwide trials* som er systemomfattende vedlikeholdsstudier der en evaluerer modeller for å vedlikeholde og overvåke kvaliteten av systemomfattende implementering over tid. Gjennom å beskrive forskningen som en syklus tenker de seg at hver fase bygger på den foregående i en logisk sekvens, samtidig som hver av fasene har sitt optimale forskningsdesign og kriterier for hva som skal gjelde som evidens. Mens antallet effekt- og effektivitetsstudier har vist en jevn økning, har det vært lite forskning på de neste fasene. En av årsakene til dette er at det til nå har vært rapportert om få prosjekter der en legger opp til

implementering i stor skala, det vil si på nasjonalt eller regionalt nivå (Kellam & Langevin, 2003). Kellam og Langevin (2003) drøfter hvilke forhold som bidrar til at programmer kan implementeres innenfor et større område. De peker på nødvendigheten av å ha sosiale og politiske strukturer som kan støtte opp om implementering i stor skala, og hva som skal til for å vedlikeholde programmene med høy grad av implementeringskvalitet. De peker særlig på behovet for strukturer for opplæring, veiledning, praktisk programstøtte og monitorering. Kellam og Langevin (2003) anbefaler at fasene skal bygge på hverandre slik at det som læres i hver fase danner grunnlag for den neste generasjonen av studier.

Glasgow, Lichtenstein og Marcus (2003) er imidlertid uenige i at det bare er de beste intervensjonene i effektstudier som skal testes i effektivitetsstudier. Den logiske progresjonen som Kellam og Langevin (2003) beskriver kan etter deres vurdering føre til satsing på intervensjoner som har liten mulighet for å lykkes ordinær praksis. De tar derfor til orde for å starte direkte med effektivitetsstudier, identifisere eventuelle modererende faktorer som begrenser intervensjonenes robusthet og bearbeide disse for å styrke intervensjonenes ytre validitet og generaliseringsmulighetene.

Implementeringsforskning er et relativt nytt forskningsområde der en blant annet har forsøkt å identifisere hva som hemmer og fremmer iverksetting og gjennomføring av programmer som har vist seg å virke i effekt- og effektivitetsstudier. Program implementering handler om i hvilken grad programmet iverksettes i samsvar med programutviklernes intensjoner. Slik forskning er særlig relevant når en skal teste om et program fungerer i ordinær praksis gjennom å vise at det er sammenheng mellom implementeringsgraden og resultatene. Mange av implementeringsstudiene bygger på Rogers (1995) teori om ”diffusion of innovation” der han beskriver hvilke forutsetninger og kjennetegn som bidrar til en vellykket spredning av nye intervensjoner. Noen av de viktigste bidragene har kommet fra enkeltstudier (Elliott & Mihalic, 2004; Mihalic & Irwin, 2003), og forskningssynteser blant annet fra Fixsen mfl. (2005), Greenberg, Domitrovich, Graczyk & Zins (2005) og Durlak & DuPre (2008)

Implementering av evidensbaserte programmer

Basert på en systematisk oversikt over cirka 500 undersøkelser beskriver Durlak og DuPre (2008) hvilke elementer som inngår i arbeidet med å overføre virksomme program til ordinær praksis. Det første handler om hvor godt informasjonen om det enkelte program formidles eller spres til lokale miljøer (eng. dissemination), det andre hvordan en lokal organisasjon

eller gruppe tar en beslutning om å prøve det nye programmet (eng. adoption), det tredje hvor bra programmet iverksettes og gjennomføres i en forsøksperiode (eng. implementation) og det fjerde elementet om hvor godt programmet opprettholdes over tid (eng. sustainability).

Programutvikling og formidling

De fleste behandlingsprogrammene for alvorlige atferdsproblemer blant barn og ungdom er utviklet i USA, og bygger i hovedsak på nord-amerikansk longitudinell og eksperimentell forskning. Ett unntak fra dette mønsteret er Olweus' (1992) anti-mobbe program som ble utviklet i Norge og senere implementert i en rekke land, og ble også innlemmet i den såkalte Blueprint serien (Elliott, 1998). Formidling representerer programutviklernes perspektiv på spredning av programmet og handler om hvor stor formidlingskapasitet de har og hvor godt de lykkes med å spre programmet til nye brukergrupper og miljøer. Rogers (1995) framhver fire forhold som fremmer formidlingen av innovasjoner i betydningen ny teknologi og ideer: (1) innovasjonens kjennetegn, (2) kommunikasjonskanalene for de nye ideene, (3) timingen dvs. tidspunktet for introduksjonen av innovasjonen og (4) kjennetegn ved det sosiale systemet som innovasjonen gjennomføres i. Det må med andre ord finnes et virksomt program som på et gunstig tidspunkt kan formidles gjennom funksjonelle kommunikasjonskanaler til mottakelige praksismiljøer som har motivasjon og forutsetninger til å implementere innovasjonen. I tillegg handler det om programutviklernes kapasitet til å lære opp, veilede og tilby praktisk eller teknisk programstøtte til nye utprøvingssteder.

Funksjonelle formidlings- eller kommunikasjonskanaler beskrives av Best, Terpstra, Moor, Riley, Norman og Glasgow (2008) med begrepene linjære modeller, relasjonsmodeller og systemmodeller. I *den linjære modellen* overføres kunnskap som en enveis prosess der forskere produserer ny kunnskap og formidler denne til brukere og beslutningstakere som så innarbeider den i praksis eller policy. Denne formen for kunnskapsformidling blir kritisert fordi den kan virke autoritær og gir liten mulighet for dialog. Enkelte har stilt spørsmål ved tendensen til det de oppfatter som en 'ovenfra og ned' holdning til utviklingen av praksisfeltet når forskerne definerer premissene for god praksis (Ferrer-Wreder, Stattin, Lorente, Tubman & Adamson, 2004). Andre påpeker imidlertid at denne prosessen også kan initieres 'nedenfra-og-opp' hvis ledere og praktikere på grunnplanet etterspør evidensbaserte programmer. De kan dermed påvirke policy i betydningen beslutninger om hvordan tjenestetilbudet skal utformes og hvilke tiltak det skal satses på (Nutley mfl., 2007).

Ut fra en ideell målsetting om at praksis skal være forskningsbasert kan en også hevde at forskningen skal være praksisrelevant og praksisutviklende. Dette ivaretas bedre i

relasjonsmodellen som beskriver hvordan kunnskap utveksles mellom forskere, beslutningstakere og praktikere basert på direkte kontakt og samhandling mellom forskning og praksis. I den siste modellen *systemmodellen* retter en også oppmerksomheten mot hvordan forskningsresultater og policy tilpasses organisatoriske og kontekstuelle forhold. Det innebærer at en tar hensyn til at forholdet mellom forskning og policy blir formet, integrert og organisert i et system som består av organisasjoners strukturer, prosesser og kontekst (Best mfl., 2008). Et tilsvarende fokus på organisatoriske og kontekstuelle forhold finnes også hos Fixsen mfl. (2005; Fixsen, Blasé, Naom & Wallace, 2009) og Glasgow mfl. (1999).

Adopsjon av programmer

Adopsjon viser til en beslutning, som regel på organisasjonsnivå, om å prøve ut et program eller å la det være. Begrepet 'adopsjon' ble introdusert av Rogers (1995) for å beskrive hvordan spredningen av en innovasjon foregår. Han skiller mellom 'early adopters', 'late adopters' og 'refusers'. De som tidlig adopterer innovasjoner omtales ofte i positive vendinger som portvakter for forandring eller 'ildsjeler' (eng gatekeepers of change, champion advocates). Avgjørelsen om å adoptere et program fungerer som et startsignal for implementeringen, men er ingen garanti for en vellykket iverksetting. Det er først og fremst beslutninger på organisasjonsnivå som avgjør om et miljø velger å prøve nye innovasjoner mens enkeltindividenes holdning betyr lite i denne fasen. Rogers (1995) identifiserte fem organisatoriske kjennetegn som påvirker sannsynligheten for at innovasjoner blir adoptert: (1) relative fordeler: hvor virksom innovasjonen er i forhold til eksisterende praksis, (2) kompatibilitet: hvor godt innovasjonen er tilpasset det eksisterende systemet, (3) kompleksitet: hvor komplisert innovasjonen er, (4) prøvbarhet: om innovasjonen kan prøves ut i liten skala og (5) observasjonsmuligheter: om det er mulig å observere andre som har tatt innovasjonen i bruk. Mye tyder på at problembevisstheten i organisasjonen og dermed dens 'readiness' for endring påvirker interessen for å adoptere ny praksis. 'Readiness' handler om hvilken etterspørsel og støtte det er for nye tiltak hos lokale ledere, praktikere og brukere. Denne bygger igjen på hvor stor bevissthet og lokal oppmerksomhet det er rundt problemer samt hvor motiverte og forpliktet praktikerne er i forhold til nye tiltak (Price, 2002). Det finnes imidlertid forskning som tyder på at 'readiness' ikke er så viktig for vellykket implementering som tidligere antatt og at andre 'implementeringsdrivere' som rekrutteringsprosedyrer, kvaliteten på opplæringsprogrammer og veiledning samt støtte fra organisasjonsledelsen, kan være mer avgjørende (Fixsen mfl., 2005).

Programimplementering og programintegritet

Det kan skilles mellom programintegritet og behandlingsintegritet, der programintegritet beskriver hvordan evidensbaserte program implementeres eller iverksettes i tråd med programutviklernes retningslinjer og i samsvar med kliniske eller andre forskningsfunn som disse bygger på. Behandlingsintegritet derimot handler om hvor godt arbeidet i den enkelte sak eller med den enkelte klient ivaretar behandlingsprogrammets hovedprinsipper og komponenter (eng. treatment adherence eller fidelity). Program- og behandlingsintegritet er en forutsetning for hverandre og skal bidra til at en unngår 'avdrift' eller 'utvanning' av evidensbaserte programmer. I dette avsnittet omtales programintegritet mens behandlingsintegritet er tema i et senere kapittel.

Fixsen mfl. (2009) identifiserte og analyserte forhold som hemmer og fremmer implementering og videreføring av evidensbaserte behandlingsprogrammer. Deres modell bygger på en syntese av forskningslitteraturen om implementering og identifiserte 7 kjernekomponenter eller 'implementeringsdrivere' som samvirker i implementeringen og vedlikehold av evidensbaserte programmer eller andre innovasjoner: (Fixsen mfl., 2005; 2009):

- *Rekruttering og utvelgelse av praktikere*: Når en innovasjon blir introdusert må nye praktikere ansettes for å jobbe med brukere i henhold til metoden,
- *Opplæring*: Innovasjoner innebærer nye måter å jobbe på og krever ofte at praktikere og deres kolleger tilegner seg ny kunnskap, ferdigheter og holdninger,
- *Veiledning/konsultasjon/coaching*: som kan omfatte personlig observasjon, instruksjon, feedback, emosjonell støtte, opplæring på jobben, og debriefing,
- *Evaluering av arbeidets kvalitet og kvantitet* dreier seg om hvordan en evaluerer praktikernes bruk av innovasjonen.
- *Data-systemer som beslutningsstøtte*: Informasjon om bruk av intervensjonen i organisasjonen og som brukes til kvalitetsforbedring, eller beslutninger og avgjørelser om bruk av intervensjonen,
- *Organisasjonsendring* med nye strukturer og prosedyrer for å støtte og oppmuntre bruken av en ny intervensjon,
- *Lederskap*: organisasjonen har ledere på forskjellige nivåer som tar avgjørelser som påvirker måten praktikerne jobber med klienter/brukere på,
- *Systemendring*: endringer i de eksterne systemene i regionen eller nasjonalt for å utvikle bedre støtte for bruken av en innovasjon (fremgangsmåter, ledelse, driftsstrukturer eller metoder),

- *Implementeringsklima:* Organisasjoner har en ”personlighet” som reflekteres i den daglige driften av organisasjonen og måten personalet ser på arbeidet sitt på.

Komponentene beskrives som integrerte og gjensidig utfyllende, det vil si at de samvirker i implementeringsprosessen, og kan i noen grad kompensere for hverandre. Fixsen mfl. (2009) bruker begrepet ’purveyors’ om individer eller grupper som arbeider for å implementere et program med høy grad av integritet og god effekt (for eksempel MST services eller Atferdssenteret). Fordelen med å ha slike formidlere er at de kan akkumulere kunnskap over tid, de kan identifisere barrierer for implementeringen og utarbeide tiltak for å forsere disse. Denne eller disse aktørene kan ha en viktig rolle gjennom å forutse mulige problemer som vil oppstå når en skal implementere et nytt program, og også hvordan disse kan løses. Dette kan være særlig viktig tidlig i implementeringsfasen hvor programmer ofte synes å være mest sårbare. Det kan dreie seg om nye stillinger som implementeringsekspert som har god kunnskap om selve programmer og om hvordan de skal implementeres eller som ’kunnskapsmeglere’ (eng. knowledge brokers) som holder organisasjonen oppdatert om utviklingen på forskningsfeltet (Fixsen mfl., 2005; Nutley mfl., 2007). Det kan også dreie seg om nye oppgaver knyttet til rekruttering, opplæring, veiledning, eller til monitorering av program- og behandlingsintegritet og produktivitet.

Som et supplement til Fixsen og kollegers ’implementeringsdrivere’ har Durlak og DuPre (2008) og Greenberg mfl., (2005) oppsummert faktorer som fremmer implementeringskvaliteten av intervensjoner i følgende liste: (1) Kontekstuelle eller lokalmiljø faktorer som systemstøtte, økonomiske ressurser, policy og ’readiness’ for forandring, (2) organisasjonsfaktorer knyttet til arbeids- eller utprøvningsstedet som organisasjonsstruktur og –kultur, blant annet lederskap, lojalitet og samhold, (3) kjennetegn ved personalet på utprøvningsstedet så som kompetanse, opplæring, teoretisk orientering, åpenhet for forandring, self-efficacy og implementeringsferdigheter, (4) kjennetegn ved intervensjonen så som relevans, brukervennlighet, forskningsbasert, kjernekomponenter, (5) leverandørorganisasjonens kapasitet og kjennetegn som positivt arbeidsklima, endringsorienterte organisasjonsnormer, integrering av nye og eksisterende tiltak, felles visjoner/verdier. Videre om felles beslutninger, koordinering med andre tjenester, åpen og effektiv kommunikasjon og gode oppgaveformuleringer, (5) kjennetegn ved personalet som leverer eller implementerer programmet (eksterne/interne endringsagenter) så som lederskap, engasjement, at implementørene er kompetente, troverdige og anvender en interaktiv formidlingsstil, og (6) teknisk støtte/assistanse som opplæringsprogram, veiledning,

monitorering av fidelity, håndbøker etc. Denne oversikten overlapper delvis med begrepsmodellen til Fixsen mfl. (2005; 2009), men utdyper forhold som bidrar til en vellykket implementering så som kontekstuelle forutsetninger, kjennetegn ved intervensjonen, leverandørorganisasjonen (eng. purveyors) inkludert teknisk assistanse og utprøvningsstedet. Disse faktorene må også betraktes som komplementære, slik at manglende forutsetninger ved utprøvningsstedet (for eksempel mangel på lokal kompetanse og miljømessig støtte) til en viss grad kan kompenseres av leverandørorganisasjonens organisatoriske og faglige kompetanse samt tekniske støtte.

I "Blueprints for Violence Prevention Initiative" evaluerte Mihalic og Irwin (2003) forhold på individ- og systemnivå som bidro til en vellykket implementering av programmer. De fant at kvaliteten av teknisk støtte, ideelle programkjennetegn, lite turnover i personalet og støtte fra lokalmiljøet var blant de mest betydningsfulle faktorene. Dette understreker betydningen av stabilitet i utprøvningsorganisasjonen og støtte i nærmiljøet for en vellykket implementering. Utfordringen blir dermed å utvikle og evaluere strategier som gjør det fordelaktig og bekreftende for organisasjoner eller tjenester å implementere og vedlikeholde evidensbaserte tiltak .

Spenningsforholdet mellom programlojalitet og tilpasning til lokale forhold

Når programmer skal implementeres i på nye steder eller i arbeid med nye målgrupper oppstår det et naturlig spenningsforhold mellom behovet for høy grad av behandlingsintegritet (eng. treatment fidelity eller adherence) og behovet for tilpasning til den lokale behandlingskonteksten. Noen anbefaler betydelige tilpasninger til lokale forhold blant annet for å unngå lokal motstand mot forandring (også omtalt som re-invention, Price, 2002). Andre advarer mot å fjerne seg så mye fra det opprinnelige programmet at en beveger seg over i "the zone of drastic mutation" (Ferrer-Wreder mfl., 2004). Spørsmålet er hvor mye en kan endre eller modifisere et program før det mister behandlingseffekten som ble funnet i originalforskningen. I evalueringen av "Blueprint" serien av program for forebygging og behandling av atferdsproblemer, fant Mihalic mfl. (2004) at behovet for lokal tilpasning i de fleste tilfeller var overvurdert, og at det i praksis forekom få slike. Fixsen mfl. (2005) viser samtidig til at de aller fleste programmer vil endres over tid, men at det er først når et program er tilfredsstillende implementert at en kan videreutvikle og eventuelt forbedre det.

'Sustainability' – vedlikehold av atferdsendring og programaktiviteter

Vedlikehold kan drøftes på individ- og miljønivå. Vedlikehold av behandlingseffekter kan studeres gjennom oppfølgingsstudier for å kartlegge om de atferdsendringene som oppnås gjennom behandlingen vedvarer. På miljø- eller systemnivå har en fokus på om selve programmet vedlikeholdes, og om praktikerne fortsetter å levere tjenester eller utfører behandling med høy grad av integritet (Kellam & Langevin, 2003). Selv om et program har blitt implementert med høy grad av integritet, så kan en ikke forutsette at det vil fortsette å fungere like bra over tid. Endringer i målgruppens kjennetegn og forutsetninger eller miljøendringer vil uvegerlig over tid føre til redusert tiltaksintegritet og mindre effektivitet. Løpende kvalitetssikring kan imidlertid bidra til at programmets integritet vedlikeholdes og at kvaliteten på behandlingen opprettholdes. Ideelt sett bør virksomme behandlingsprogrammer over tid påvirke prevalensen av problemer i befolkningen. Gjennom å etablere systemer for monitorering kan en løpende følge med hvordan det går med målgruppen eller hvordan populasjonen fungerer eller har det over tid (Biglan & Ogden, 2008). Når det gjelder barn og unge med alvorlige atferdsproblemer kan en for eksempel følge utviklingen på kriminalitetsstatistikken, omfanget av plasseringer utenfor hjemmet og antall henvisninger til psykisk helsevern eller barnevernstjenesten.

Programutviklernes og evaluatorenes preferanse

Siden det er en formidabel oppgave å utvikle, teste, formidle og videreutvikle program er det viktig å spørre hvem som bestemmer hvilke program som skal implementeres og evalueres. I utgangspunktet må noen være tilstrekkelig motivert og dessuten ha økonomiske og faglige ressurser til å teste nye tiltak. Utviklingen har gått i retning av at programutvikleren og evaluatoren ofte er én og samme person, noe som har ført til spørsmålet om forskerens uavhengighet. Rollen som evaluatør har endret seg fra å være en uavhengig aktør til en utvidet rolle som ofte omfatter planlegging, gjennomføring, evaluering og formidling av program (Lipsey, 2000). Evalueringsstudier som har blitt utført av modellutviklerne, deres studenter og kolleger, har derfor av noen blitt kritisert for å ha preferanse-effekter (eng allegiance) i resultatene (Wampold, 2001). Wampold (2001) hevder at preferanser for et program eller en modell konsekvent fører til høyere effektstørrelser og det påpekes at forskere med preferanser ofte sammenligner sine programmer eller modeller med andre former for behandling som ikke er like verdsatt.

Forskningsteamenes ekspertise og preferanser har vist seg å påvirke resultatene i behandlingsforskningen og behandlingen som favoriseres av forskeren har en tendens til å gi

de beste resultatene (Luborsky mfl, 1999). Enhver terapi har en tendens til å komme bedre ut i sammenligning med andre intervensjoner når den gjennomføres av personer som er kunnskapsrike eksperter. Konklusjoner fra flere meta-analyser og kunnskapsoversikter bekrefter at involvering av evaluatøren i design og implementering av en intervensjon påvirker resultatene og effektstørrelsene i positiv retning (Petrosino, 2004). Chambless og Hollon (1998) mener at det dreier seg om ærlige forskjeller i kapasiteten til en hvilken som helst forskergruppe til å lykkes med implementeringen heller enn om et ondsinnet forsøk på å påvirke resultatene. Tilsvarende hevder Lipsey (1995): "A cynical view might attribute this to some biasing or 'wish-fulfilling' influence researchers have on the outcome of studies they control. I see another interpretation as plausible... When a researcher is closely involved in treatment design... there is likely to be a high level of treatment integrity".

Selv om kontrollbetingelsene er strukturerte med tydelige innholdsbeskrivelser, så kan utøverne være mindre entusiastiske og arbeide i et mindre støttende miljø enn det som eksperiment terapeutene jobber i (Sprenkle & Blow, 2004). For å unngå slike preferanse-effekter foreslår derfor Crits-Christoph mfl. (1991) at behandlingsforskning bør utføres av 'multiallegiance' forskningsteam, dvs team som har preferanser i forhold til mer enn ett program eller én modell. For å unngå påstander om preferanse effekter er også replikasjoner viktige, og helst replikasjoner som utføres av uavhengige forskerteam. Dette beskytter i noen grad mot kritikk om 'forsker-bias' eller at en bygger på funn som er unike for en bestemt setting eller gruppe av terapeuter (Chambless & Hollon, 1998).

Oppsummerende diskusjon

Evidensbasert policy handler om hvordan forskning legger premisser for beslutninger og prioriteringer i politikk og forvaltning og hvordan dette påvirker og påvirkes av praksis. Delvis kan det forklares som den større sammenhengen som utviklingen av evidensbaserte behandlingsprogrammer inngår i, men det har også til dels vært et konkurrerende perspektiv kjennetegnet av et mindre forpliktende forhold mellom forskning og praksis. Prinsippet om å bygge tiltak og tjenester på den best tilgjengelige forskningskunnskapen fikk i en periode stor aktualitet i Storbritannia der evidensbasert policy ble forbundet med Blair regjeringens forsøk på å reformere og modernisere statlig politikk. 'New public management' tok utgangspunkt i at praksis skulle være evidensbasert og bli grundig evaluert, men kom også til å bli forbundet med overdreven byråkratisk styring og kontroll med praksis. Denne sammenkoblingen kan

antagelig være med på å forklare hvorfor evidensbasert praksis har blitt omstridt i norsk sammenheng (bl.a. Fossestøl & Enehaug, 2008; Skulberg, 2008).

Det finnes med andre ord to hovedlinjer i utviklingen av evidensbasert praksis. Den ene har sitt utspring i evidensbasert medisin (Sackett mfl., 1996) og har blant annet påvirket utviklingen av evidensbasert psykologisk praksis (APA, 2006; Rønnestad, 2008) og kunnskapsbasert sosialt arbeid (Socialstyrelsen, 2000; Jergeby & Sundell, 2008). Den andre utviklingslinjen som fokuseres i denne avhandlingen er knyttet til evidensbaserte programmer som er forhåndsdefinerte, strukturerte og til en viss grad standardiserte intervensjoner (APA, 2006; Elliott, 1998). Implementeringen MST og PMTO i Norge startet som et direkte og konsensusbasert samarbeid mellom forvaltning og forskning der samarbeidet bygde på en felles forståelse av at alvorlige atferdsproblemer bør ha høy prioritet, at internering og gruppetiltak ikke hadde gitt så gode resultater som forventet og at familie- og nærmiljøbaserte modellprogram kunne føre til bedre resultater, økt tilgjengelighet og høyere behandlingsproduktivitet.

Begrepsmodellene som illustrerer elementene som enkeltvis og samlet bidrar til en vellykket implementering av ny praksis har mange likhetstrekk. De viser først og fremst at mange faktorer virker inn og mye må stemme før en lykkes med å implementere nye programmer eller tiltak. Hvis et intervensjonsprogram ikke gir positive resultater kan forklaringen være at programmet ikke er virksomt, men også at det er mangelfullt eller dårlig implementert. Et program kan være forsinket, ufullstendig eller feilaktig implementert, noe som fører til at noen klienter faktisk ikke får det tilbudet eller de tjenestene som programmet lover. Det kan dreie seg om program som ikke har rukket å bli riktig implementert, eller om program som aldri blir skikkelig installert eller etablert. Men det kan også dreie seg om programmer som har vært i drift over lengre tid og som etter hvert har problemer med å vedlikeholde høy grad av integritet. Disse forholdene begrunner hvorfor det er behov for fokus på og forskning om selve implementeringsprosessen, det vil si på spredningen av virksomme program til nye utprøvningssteder. Over tid vil implementeringen av evidensbaserte programmer ha organisatoriske konsekvenser og stille nye krav til organisering og ledelse av tjenester eller virksomheter.

Storbritannia har vært den fremste eksponenten for policyinformert praksis i tråd med det brede perspektivet på policy i tråd med opplysningsmodellen og med det som kan kalles konseptuell bruk av forskningen. Det evidensbaserte perspektivet som også omtales som det smale eller instrumentelle perspektivet har vært vanligere i USA der forskning har lagt

grunnlag for modell programmer som så har blitt testet i randomiserte evalueringsstudier. Utviklingen i Norge kan sies å være påvirket både av USA og Storbritannia, selv om de evidensbaserte programmene fra USA tydeligst har preget etableringen av nye tiltak for barn og unge med atferdsproblemer (Ogden mfl., 2010). Men uavhengig av hvilket perspektiv en har på policy så ser det ut til at beslutninger om praksis sjelden er en rasjonell beslutningsprosess der tyngden i forskningsdata er avgjørende.

Synet på og holdningene til utvikling og formidling av forskningsbaserte behandlingsprogrammer har vært forskjellig i USA og Europa. Bullock (1992) hevder at det kontinentale Europa tradisjonelt har hatt en filosofisk og logisk heller enn empirisk tilnærming til løsning av ungdomsproblemer. Storbritannia derimot har en lang tradisjon for å knytte empirisk forskning og policy mens demonstrasjonsprosjekter og evalueringsstudier som 'selger' programmer passer godt til entreprenørkulturen i USA (Bullock, 1992). Både universiteter og frittstående forskningsinstitusjoner har en lang tradisjon for å omsette kliniske forskningsfunn til praksis. Flere av institusjonene finansierer forsknings- og formidlingsvirksomheten gjennom å konkurrere om forskningsmidler og gjennom å selge programlisenser og 'know-how' til brukermiljøer. Denne blandingen av forskning og formidling har legitimitet i de akademiske miljøene, og har ført til at USA er en dominerende internasjonal leverandør av behandlingsprogrammer. For norske og europeiske forskere kan det være få insentiver og lite som motiverer til formidling av forskningsresultater til praksisfeltet. Det kan være vanskelig å få avsatt tid til formidling og i de akademiske institusjonene finnes det kulturelle så vel som praktiske barrierer for slik virksomhet (Ferrer-Wreder mfl., 2004; Nutley mfl., 2007).

Utviklingen av evidensbasert praksis i Norge har vist at de fleste forebyggende eller tiltaksrettede programmene har blitt utviklet eller oversatt og implementert basert på internasjonal forskning, særlig fra USA. De norske evalueringene kom i ettertid, som replikasjoner av studier som programutviklerne har gjennomført. Det gjelder også for behandlingsprogrammene MST og PMTO som ble besluttet implementert på nasjonal basis før det forelå klare forskningsmessige indikasjoner på at de fungerte i norsk sammenheng. Kellam og Langevins (2003) fasemodell som forutsatte at effektivitetsforskningen skulle komme før effektivitetsforskning og deretter følges av implementering i stor skala ble ikke fulgt i Norge. Implementering i stor skala ble etterfulgt av effektivitetsforskning. Det dreide seg om en kombinasjon av strategisk og konseptuell bruk av forskning, heller enn instrumentell bruk. De amerikanske forskningsresultatene ble benyttet som legitimering av programmenes empiriske basis, og uten at noen på forhånd problematiserte deres overførbarhet til norske forhold. Den

norske effektivitetsforskningen foregikk deretter parallelt med implementeringen, og resultatene fikk først og fremst en bekreftende funksjon gjennom å vise at programmene virket.

3. Evidensbaserte behandlingsprogram – programfaktorer, fellesfaktorer og klinisk skjønn

Evidensbaserte behandlingsprogram blir her brukt om Evidence-Supported-Treatment (EST) som Chambless og Hollon (1998) definerer som : “... a clearly specified psychological treatments shown to be efficacious in controlled research with a delineated population” (s. 7). Det dreier seg med andre ord om en definert behandling som har vist seg virksom (eller effektiv) for bestemte målgrupper gjennom kontrollert evalueringsforskning. I dette kapitlet presenteres og drøftes evidensbaserte behandlingsprogrammer, blant annet i forhold til hvordan resultatene påvirkes av behandlingsprotokoller, behandlingsintegritet, tilleggsproblemer (komorbiditet) og behandlingens lengde. Videre drøftes forholdet mellom programfaktorer og terapeutiske fellesfaktorer og forholdet mellom standardisert behandling og klinisk skjønn. Blant fellesfaktorene drøftes betydningen av terapeutenes kompetanse og klientenes kjennetegn for behandlingsutfall og hvordan den terapeutiske alliansen kan påvirke resultatene. Selv om hovedfokuset i denne sammenhengen er rettet mot psykologisk behandling av atferdsproblemer hos barn og unge, blir resultater og erfaringer fra andre behandlingsområder og aldersgrupper trukket inn der det er relevant for sammenhengen.

Kommentar om begrepsbruk

Evidensbaserte behandlingsprogrammer blir i teksten brukt om det som på engelsk omtales som “empirically supported therapy or treatment”. Over tid har den engelskspråklige begrepsbruken endret seg fra *evidence-based programs* (evidensbaserte program) til *empirically validated treatments* (empirisk validert behandling) og senere igjen til *empirically supported treatments* (emprisk støttet behandling) (Chambless & Ollendick, 2001). Selv om begrepsbruken i fag- og forskningslitteraturen ikke er konsekvent, så kan det se ut til at evidensbegrepet er noe neddempet, muligens fordi det har vært kontroversielt, og for å understreke at det dreier seg om forskning/empiri mer enn om ‘bevis’. Det er imidlertid ikke uvanlig at en i publikasjoner mer generelt omtaler “evidence-based treatment” i betydningen evidensbasert behandling.

Evidensbasert praksis er et mer omfattende begrep en evidensbaserte programmer og handler om metoder, program, modeller eller arbeidsmåter som kan samles under begrepet ‘intervensjoner’ eller ‘tiltak’. Den amerikanske psykologforeningen (APA, 2006) introduserte

'evidence based psychological practice' (EBPP) for å understreke at forskningen ikke bare skulle fokusere på intervensjonenes effektivitet, men også på deres anvendbarhet, nytte og generaliserbarhet. Det ble også understreket at praksis skulle ta hensyn til klientenes kjennetegn, kultur og preferanser og denne forståelsen ligger nær opp til Sackett og kollegers (1996) definisjon av evidensbasert medisin.

Intervensjoner kan være metoder, program, retningslinjer, strategier eller teknikker. Med andre ord beskriver intervensjoner tiltak som formidles i tjenester for barn og unge (helse, sosial, skole) og kan omfatte kartlegging/ utredning, diagnosistering, behandling, psykoterapi og ettervern. Og mens *metoder* beskriver behandlingsmodaliteter (for eksempel kognitiv atferdsmodifikasjon) så handler *program* om behandling (eller forebygging) som består av flere komponenter.

Behandling. Behandlings- og terapibegrepet er uklart i arbeidet med atferdsproblemer blant barn og ungdom, og det er ulike oppfatninger av hva som bør rubriseres som behandling eller terapi til forskjell fra andre intervensjoner. Uklarheten henger blant annet sammen med at utviklingen av tiltak på dette feltet i USA i stor grad har foregått innenfor barne- og ungdomspsykiatrien hvor terapi (for eksempel Multisystemisk Terapi) og behandling er vanlige begreper. I Norge har arbeidet med atferdsproblemer i stor grad også vært knyttet til skoleverket (blant annet spesialskoler og verneskoler) og barnevernet, der disse begrepene har vært mindre vanlige. Behandling og terapi blir derfor brukt i vid betydning i denne teksten for å beskrive intervensjoner som retter seg mot barn og unge som har alvorlige atferdsproblemer i klinisk forstand. Psykoterapi brukes når det framgår av originallitteraturen at det mer avgrenset dreier seg om behandling utført av kvalifiserte psykologer eller psykiatere. Et annet begrep som kan virke fremmedartet er treningsbegrepet, for eksempel i Parent Management Training, og som dekker både behandlings- og opplæringstiltak for foreldre (eventuelt for barn som i Social Skills Training). Videre kan skillet mellom forebyggende program og behandlingsprogram være uklart på dette feltet fordi behandling av atferdsproblemer hos barn og ungdom også kan betraktes som forebygging av videre problemutvikling (Weisz, Sandler, Durlak & Anton, 2005).

Klienter brukes i denne sammenhengen heller enn pasienter og omfatter barn, unge, foreldre og familier og kan også omtales som brukere eller målgruppe.

Randomiserte kontrollerte forsøk (randomised controlled trials – RCT) er det vanligste forskningsdesignet i intervensjonsforskningen og omtales også som eksperimenter. Det kan imidlertid virke teknisk og fremmedgjørende når en i studier som omfatter individer snakker om eksperimenter, utvalg, eksperiment- og kontrollgrupper, så disse begrepene blir i de

følgende kapitlene ofte erstattet med henholdsvis randomiserte kontrollerte undersøkelser, deltakere, intervensjonsgruppe og sammenligningsgruppe.

Som hovedregel blir '*effektforskning*' (eng. efficacy research) brukt om testing av programmer under forskningsoptimale forhold mens '*effektivitetsforskning*' (eng. effectiveness research) blir brukt om forskningsbasert testing av programmer i ordinær praksis. Når terapeuter arbeider i tråd med retningslinjene for evidensbaserte programmer omtales det som *behandlings-integritet* eller behandlingslojalitet. Behandlingsintegritet brukes med andre ord om 'treatment integrity', men også om 'treatment fidelity' og 'treatment adherence'.

Av og til har det vært vanskelig å oversette engelske begreper til norsk slik at en ivaretar deres begrepsmessige presisjon eller nyanser. Derfor blir de engelske begrepene av og til føyd til de norske i parentes.

Effekt og effektivitetsstudier

Randomiserte studier som *effektstudier* kjennetegnes av homogene deltakergrupper og at forskeren har stor grad av kontroll med gjennomføringen (omtales også som "demonstration projects" eller "research therapy projects"). Effektstudier gir muligheter for å si noe om kausale sammenhenger mellom intervensjon og resultater og forutsetter at eksperimentgruppens observerte utbytte er et resultat av selve behandlingen og ikke av tilfeldigheter, forløp av tid, effekt av kartlegging, eller av at de ulike behandlingsbetingelsene har forskjellige deltakere (Chambless & Hollon, 1998). En begrensning ved slike studier er imidlertid at de ikke fanger opp om intervensjonen fungerer i ordinær praksis med varierte klientgrupper. I forskningslitteraturen tas det derfor ofte til orde for å styrke innslaget av effektivitetsstudier som et supplement (Brestan & Eyberg, 1998; Kazdin, 2008; Kellam & Langevin, 2003) eller som et alternativ (Glasgow mfl., 2003) til effektstudier.

Effektivitetsstudier med randomiserte gruppedesign innebærer som regel at deltakergruppen er mer heterogen og at intervensjoner evalueres i ordinær praksis dvs under betingelser som ligner på eller tilsvarer de som tiltaket senere skal praktiseres i. Forskeren har som regel mindre kontroll over rammebetingelsene og den praktiske gjennomføringen av behandlingen i slike studier, og de som skal praktisere programmet får vanligvis mindre opplæring, veiledning og støtte enn praktikerne som bidro til utviklingen av den første prototypen. Effektivitetsstudier omtales også som "real-life projects". Etter at et program er testet i effektstudier kan det neste trinnet være å teste det under ordinære praksisbetingelser,

enten som et ledd den forskningsbaserte utprøvingen av programmet eller som et ledd i prosessen der en overfører virksomme programmer fra forskning til praksis. Videre vil en kontrollere at opplæring og veiledning har tilstrekkelig kvalitet og at behandlingen gir like positive behandlingsresultater som forventet, basert på effektstudiene.

Glasgow mfl. (2003) sammenligner effektstudier med effektivitetsstudier på følgende måte: Effektstudier har homogene deltakergrupper med høyt motiverte individer mens effektivitetsstudier har et bredt sammensatt heterogent utvalg, ofte i form av en definert populasjon. Effektstudier har intensive, spesialiserte og standardiserte intervensjoner med utvalgte terapeuter og som testes i randomiserte design der en forsøker å maksimere effektstørrelsen mens effektivitetsstudier har intervensjoner som ikke krever stor ekspertise og som også kan evalueres i randomiserte design, eller med tidsserie- eller kvasi-eksperimentelle design. I effektstudier testes intervensjonen med bruk av eksperter og store ressurser, vanligvis i én setting for å redusere variasjon mens effektivitetsstudier evalueres i flere settinger og kan i noen grad tilpasses lokale forhold. I effektstudier implementeres behandlingen av en forskningsstab som nøye følger en protokoll mens i effektivitetsstudier implementeres intervensjonen med bruk av en modifisert protokoll og av varierende personalgrupper som kan ha konkurrerende oppgaver.

Det særlig på to områder at effektivitetsstudier skiller seg fra effektstudier. For det første, mens effektstudier har en homogen og veldefinert målgruppe samt standardisere utprøvningssteder, så har effektivitetsstudier deltakere som kan variere mye med hensyn til sosial bakgrunn, problembelastning, risikonivå og tilleggsproblemer. For det andre kan praktikerne eller behandlerne i effektivitetsstudier ha varierende utdanningsnivå og tjenestetilknytning

. Chambless & Ollendick (2001) identifiserte 118 manualiserte behandlingsformer og fant at effektivitetsstudier hadde høyere ytre, men lavere indre validitet enn effektstudiene. Det vil si at effektstudiene var bedre egnet til å fastslå at det var intervensjonen og ikke andre forhold som forårsaket resultatene mens effektivitetsstudiene ga resultater som hadde større overføringsverdi til praksis. Det viste seg også at klientene i effektivitetsstudier verken forbedret seg like mye eller oppnådde samme positive funksjonsnivå ved avslutningen av behandlingen som klientene i effektstudier. Videre viste det seg at klientene i noen effektivitetsstudier mottok mer behandling i form av lengre behandlingsforløp eller tilleggsbehandling og andre tjenester enn det som var vanlig i effektstudier (Chambless & Ollendick, 2001). Dette kan tyde på at effektstudier gir bedre

resultater enn effektivitets-studier og at innholdet i behandlingen i mindre grad avviker fra behandlingsprotokollen samtidig som effektivitetsstudiene har større generaliseringsverdi.

Samtidig har det skjedd en tilnærming mellom de to undersøkelsesmåtene slik at effektivitetsstudiene har innarbeidet mange av forskningskravene fra effektstudiene for å kunne underbygge at det er en kausal sammenheng mellom intervensjonen og de kliniske behandlingsresultatene, samtidig som resultatene i større grad kan generaliseres til ordinær praksis (se for eksempel Curtis, Ronan & Borduin, 2004; Ogden & Amlund-Hagen, 2006; 2008).

Kritikk av programbasert behandling

Evidensbaserte behandlingsprogrammer bygger på og evalueres gjennom forskning. Det har vært en sterk vekst i antall programmer for forebygging og behandling av psykososiale problemer, og en tilsvarende økning i antall evalueringsstudier. Utviklingen av evidensbaserte behandlingsprogrammer der en anvender kliniske forskningsfunn i praksis har i følge Kazdin (2008) skjerpet motsetningsforholdet mellom forskning og praksis i klinisk psykologi. Selv om resultatene fra behandlingseffektforskningen har vært oppmuntrende, har de blitt kritisert både av klinikere og forskere. I de neste avsnittene drøftes noen av de vanligste innvendingene mot evidensbaserte behandlings-programmer slik de har blitt presentert i forskningslitteraturen (APA presidential task force on evidence-based practice, 2006; Chambless & Ollendick, 2001; Weisz, Jensen-Doss og Hawley, 2005, 2006), men er særlig utdypet i artikler av Westen, Novotny og Thompson-Brenner (2004, 2005). Kritikken av evidensbasert behandling og behandlingsforskning har blant annet blitt imøtegått av Chambless & Ollendick (2001), Kazdin (2008), Sexton, Ridley og Kleiner (2004), Weisz, Jensen-Doss & Hawley (2005; 2006), Weisz, mfl. (2005) og Weisz, Weersing & Henggeler (2005).

De viktigste innvendingene mot evidensbaserte behandlingsprogrammer er at: (1) de har et eksklusivt fokus på korttidsbehandling, (2) det er usikkert hvor anvendelige de er i forhold til klientvariasjoner i komorbiditet, personlighet, etnisitet og kultur, (3) de har blitt utviklet og testet på enkelt- problemer og diagnoser og på relativt enkle, ofte subkliniske saker og virker kanskje ikke så bra i kompliserte og alvorlige saker som man støter på i vanlig klinisk praksis, (4) de er så protokoll- eller manualdrevne at de gjør det vanskelig å individualisere behandlingen og begrenser terapeutens kreativitet når denne skal forholde seg til uvanlige eller uventede hendelser i klienters liv og (5) de mangler spontanitet og

fleksibilitet slik at de interfererer med kontaktetablering og utviklingen av en god terapeutisk relasjon.

Den vanligste innvendingen mot programbasert behandling er at den er lite fleksible og underkjenner klinikernes utdanning, erfaringsbaserte kunnskap og skjønn (Westen mfl., 2004). Praktikere kan oppleve at deres utdanning og opparbeidete kompetanse blir vurdert som mindre viktig og at programmene redusere mulighetene deres for fritt metodevalg.

Kritikken mot evidensbaserte behandlingsprogrammer har i USA har til dels vært preget av sterkt negative karakteristikk. Noen peker på at klinikere i USA har vist en tydelig mangel på entusiasme for 'empirically supported therapies' (APA, 2006, Westen mfl., 2005) og beskriver litteraturen om evidensbaserte tiltak som en "idiotsikkert, trinnvis instruksjon til paraprofesjonelle". Andre hevder at terapeuter skal være "... a disciplined improvisational artist, not a manual-driven technician" og omtaler evidensbasert behandling som "promoting a cookbook mentality" eller "painting by numbers" (Chambless & Ollendick, 2001). Westen mfl. (2005) hevder at forskernes manglende interesse for erfarne klinikers kompetanse signaliserer at forskere vil forbeholde seg retten til å tenke, lese bredt og integrere data mens klinikerne skal holde seg til boka og unngå kliniske vurderinger. De mener at hvis forskerne skal avgjøre hvilke intervensjoner klinikerne skal bruke eller ekskludere samt lengden på behandlingen, er ikke dette noe mer eller mindre "klinisk" i Meehl's forstand (1954) (dvs. subjektiv og uformell) enn lignende beslutninger som tas av klinikere. Med andre ord hevder de at forskeren forsøker å overta klinikerens rolle og påberoper seg frihet til å foreta vurderinger som de mener at klinikerne må avstå fra (Westen mfl., 2005).

Den evidensbaserte tilnærmingen har satt nye standarder for hvordan behandling skal beskrives, gjennomføres og evalueres, og utfordrer etablert praksis på flere områder. Som det framgår av de neste avsnittene kan noe av kritikken tilbakevises fordi den bygger på feil forutsetninger eller manglende kunnskap om hva evidensbaserte programmer er. Men det finnes også konstruktiv kritikk som problematiserer eller påpeker begrensninger ved evidensbaserte intervensjoner og forskningen som de bygger på. Noe av motstanden mot den evidensbaserte tilnærmingen kan forklares med at økt satsing på studier av programeffekter har ført til en nedprioritering av andre temaer i behandlingsforskningen og noen er antagelig misfornøyde med at deres praksismodeller kommer dårlig ut i sammenligninger eller stilles i skyggen av nye tilnærminger.

Noe av kritikken av evidensbaserte programmer har vært rettet mot å definere behandlingens innhold og prosedyrer gjennom utarbeiding av behandlingsprotokoller eller manualer. Dette er tema for neste avsnitt.

Behandlingsprotokoller – behandlingsmanualer i forskning og praksis

For å kunne hevde at en behandling virker er det nødvendig å beskrive hva den inneholder. Brede overskrifter som atferdsmodifikasjon eller kognitiv behandling kan dekke over store variasjoner i innhold og prosedyrer. Kunnskap om behandlingens innhold er en forutsetning dersom forskere skal kunne replikere en definert behandling, men også hvis forskningen skal ha informasjonsverdi for praktikere. For å formidle hvilken behandling de har evaluert, må forskere eksplisitt kunne beskrive hvilke teknikker og strategier den består av (Chambless & Hollon, 1998). I evidensbasert behandling blir derfor programmets viktigste komponenter, prinsipper og prosedyrer beskrevet i en behandlingsprotokoll, manual eller håndbok. Det innebærer at behandlingsprogrammer er standardiserte i den forstand at de viktigste innholdskomponentene er eksplisitte og forhåndsdefinerte. Presisjonsnivået i beskrivelsen av innholdet kan imidlertid variere med hva slags behandling det dreier seg om. En behandlingsprotokoll kan beskrive intervensjonen sesjon for sesjon, eller beskrive brede prinsipper og faser i behandlingen med eksempler på tiltak som samsvarer med disse. Retningslinjene kan være detaljerte eller de kan legge opp til utstrakt fleksibilitet innenfor programmets generelle rammer (Kendall, Chu, Gifford, Hayes & Nauta, 1998). På pluss siden kan protokoller ha en positiv funksjon gjennom å virke standardiserende, fremme terapeutenes metodekompetanse, og styrke behandlingsintegritet gjennom at terapeutene får standardisert opplæring og systematisk veiledning i bestemte modeller (Lambert & Ogles, 2004).

Behandlingsprotokoller er et viktig innslag i randomiserte kliniske studier av intervensjoner og kan brukes for å lære opp terapeuter som deltar i kliniske forsøk gjennom å konkretisere innholdet i behandlingen (Carroll & Rounsaville, 2008). De spesifiserer og operasjonaliserer den uavhengige variabelen (dvs. behandlingen) og legger grunnlag for evaluering av relevans og effekt samt for replikasjonsstudier. Videre definerer de grensene og innholdet i behandlingen slik at en kan evaluere hvor anvendelig og virksom den er. Beskrivelsen kan også omfatte hvordan behandlingen ligner på eller er forskjellig fra kontrollbetingelsene. Protokoller kan videre beskrive teoretiske begrunnelser for behandlingen, definere problemet som behandlingen retter seg mot samt presentere målene for behandlingen. Her inngår også behandlingens struktur, så som varighet, format (gruppe vs. individuell), intensitet (antall og lengdene på sesjonene), grad av fleksibilitet vs. struktur, og innholdet i sesjonene.

Behandlingsprotokoller, manualer eller håndbøker er med andre ord nærmest en forutsetning for behandlingseffektforskning, gjennom å konkretisere behandlingens innhold og hvordan den skal implementeres (Carroll & Rounsaville, 2008).

Selv om kritikerne aksepterer at behandlingsprotokoller er nødvendige i forskningen for å sikre en standardisert behandling på tvers av utprøvningssteder og klargjøre for brukerne hva som egentlig testes så ser de også ulemper for forskningen (Westen mfl., 2004). Hvis klinikerne i effektstudier skal bestemme hvordan og når en skal intervenere ut fra behandlingsprinsippene, så reduseres den eksperimentelle kontrollen. Avvik fra manualen gjør det vanskelig å trekke årsakskonklusjoner om eksperimentelle effekter og derfor må en oppgi hvor mange sesjoner som skal gjennomføres samt hva som skal skje i hver sesjon eller i sekvensen av sesjoner (Kazdin, 2008; Westen mfl., 2004). Og hvis en er opptatt av endringmekanismer eller å studere forholdet mellom prosess og resultater er behandlingsvariasjon viktig, noe som motvirkes gjennom restriksjonene som manualisert behandling legger på behandlingen. Dessuten vil rekruttering, seleksjon og innlemming av saker i forskningen avvike vesentlig fra prosessen som bringer individer til behandling i det virkelige liv. Blant annet må deltakerne i forskningsprosjekter gi sitt informerte samtykke der de akseptere at de skal randomiseres. Oppsummert vil eksperimentelle effektstudier gjennom bruk av behandlingsprotokoller legge til rette for standardiserte behandlingsforløp med gode muligheter for å påvise kausale sammenhenger mellom intervensjoner og resultater, mens variasjoner og avvik fra protokollen forebygges, fordi de kan svekke konklusjoner om årsakssammenhenger. Som tidligere nevnt vil ikke disse innvendingene være like relevante i effektivitetsstudier som er mer representative for ordinær praksis og der kontrollbetingelsene er mer moderate.

Mens få stiller spørsmål ved verdien av å bruke behandlingsprotokoller i behandlingseffektstudier, så er motstanden større når det gjelder å bruke foreskrivende protokoller i klinisk praksis (Duncan & Miller, 2005). Westen mfl. (2004) hevder at manualer reduserer klinikerens til en forskningsassistent som behandler klienter på en standardisert måte. Etter deres oppfatning skjer det når manualer går over fra å være deskriptive (hvor forskerne beskriver sine eksperimentelle manipulasjoner på en presis måte) til å bli foreskrivende (som beskriver standardisert klinisk aktivitet i vanlig praksis). Dette åpner for at psykoterapi kan bli en jobb for para-profesjonelle som uten faglige kvalifikasjoner verken kan eller skal utøve klinisk skjønn i valg av intervensjoner eller i fortolkning av kliniske observasjonsdata. Dette støttes av Duncan og Miller (2005) som hevder at manualisert behandling verken bidrar til å gi

oversikt over psykoterapifeltet eller til bedre resultater i psykoterapi. De mener at manualer vektlegger spesifikke tekniske operasjoner og beskriver psykologiske intervensjoner som om de var uavhengige av hvem som utfører dem og hvem som mottar dem. I en spissformulert kritikk hevder de at det er en viss forførende apell i å gjøre psykologiske intervensjoner idiotsikre, der brukerne – klienter og terapeuter – i hovedsak er irrelevante. Terapeuter som følger boken, utvikler ifølge Duncan og Miller (2005), en bedre relasjon til manualen enn til klienten.

Kritikken imøtegås av Weisz, Weersing & Henggeler (2005) som peker på at protokoller verken minimaliserer klientenes aktive deltakelse i behandlingsprosessen eller hindrer terapeuten i å foreta kliniske vurderinger. Behandlingsmanualer beskriver som regel intervensjonene i en logisk sekvens, for eksempel slik at kartlegging og alliansebygging kommer før intervensjonene. Videre formidles som regel prinsipper eller retningslinjer for praksis som anbefalte teknikker heller enn påbud. De kan også formidle at visse prosedyrer kan være ineffektive eller kontraproduktive. Det gjelder for eksempel når Multisystemisk Terapi (MST) og Multidimensional Treatment Foster Care (MTFC) fraråder bruk gruppeterapi i behandlingen av alvorlige atferdsproblemer fordi det kan føre til 'avvikstrening' blant klientene. Men innenfor rammene av manualenes logiske og empiriske retningslinjer har klinikerne betydelig rom for fleksibilitet og mandat til å bruke sin egen kreativitet og ressurser for å fremme atferdsendring. Så selv om det kan finnes eksempler på rigide manualer så er mange så romslige at terapeuten, klienten og andre involverte spiller en aktiv rolle i utformingen av behandlingens innhold. De fleste behandlingsmanualer for behandling av alvorlige atferdsproblemer kommer i denne kategorien (se for eksempel Henggeler mfl., 1998). Weisz, Weersing og Henggeler (2005) omtaler evidensbasert behandling som en transaksjonsprosess der terapeuter og klienter samarbeider for å konkretisere behandlingsmål og å lage planer for gjennomføringen. Samtidig erkjenner de at kompetent praktisering av evidensbasert psykoterapi handler om noe mer enn mestring av spesifikke manualbaserte prosedyrer. Variasjoner i behandlingseffekt kan skyldes forskjeller i terapeutenes profesjonelle kliniske kompetanse eller generelle terapeutiske ferdigheter og derfor må protokollene suppleres med programspesifikk opplæring og veiledning av terapeutene .

Når terapeuter oppfatter behandlingsmanualer som terapeutiske tvangstrøyer kan det skyldes at de mener at det legges for stor vekt på teknikker og at dette påvirker den terapeutiske alliansen negativt. Klinikervennlige manualer bør ifølge Carroll og Rounsaville (2008) spesifisere det som er felles for flere terapier og beskrive hvordan en kan styrke disse

fellesfaktorene og avveie dem i forhold til teknikker som er spesifikke for det enkelte program (se også Addis, Wade & Hatgis, 1999). De forventer at terapeuter vil være mer aksepterende hvis de mener at behandlingsmanualen øker, heller enn reduserer deres kliniske ekspertise (Onken, Blaine, & Battjes, 1997). Terapeutvennlige manualer bør gi eksplisitte råd om hvordan en takler klienter som er utfordrende, lite motivert, ambivalente, som viser motstand, som har vanskelig for å ordlegge seg eller som er kognitivt lavtfungerende (Carroll & Nuro, 2002).

Behandlingsprotokoller har en sentral plass i evidensbasert behandling gjennom å beskrive intervensjonen man primært er opptatt av å evaluere, men som det framgår av de foregående avsnittene er de mindre kontroversielle i klinisk forskning enn i praksis. Og mens kritikerne er mest opptatt av restriksjonene som protokollene legger på kreativitet og skjønn i klinisk praksis så er tilhengerne opptatt av å framheve mulighetene som terapeutene har til å individualisere behandlingen og som klientene har til å påvirke behandlingsforløpet innenfor protokollens definerte rammer.

Behandlingsintegritet og kompetanse

Høy grad av behandlingsintegritet (fidelity / adherence) handler om å gjennomføre behandlingen eller intervensjonen som forutsatt av programutvikleren, i tråd med mål, retningslinjer og underliggende teori. Det kan handle om hvorvidt behandlingen blir gjennomført som planlagt og intendert, men også om kvaliteten av behandlingen slik klienten opplever den. Forskingen om behandlingsintegritet er ikke konsistent og mens noen har funnet at høy grad av behandlingsintegritet er forbundet med gode resultater (Frank & Frank, 1991) så har andre funnet at høy eller for høy integritet kan gi dårligere resultater (Castonguay, 1996; Sánchez mfl., 2007).

Sammenhengen mellom integritet og resultater kan i noen tilfeller være ikke-linjære noe som understreker viktigheten av å undersøke kurvilinjære sammenhenger. Det kan også dreie seg om takeffekter der høy integritet gir bedre resultater inntil et visst punkt. Målingene av slike sammenhenger er imidlertid omdiskutert for behandlingsintegritet ofte måles som korrelasjoner, noe som kan gjøre det vanskelig å fastslå årsaksretningen; om høy integritet førte til gode resultater eller om et positivt forløp i terapien gjorde det lettere for terapeuten å følge behandlingsprogrammet (Webb, DeRubeis & Barber, 2010).

Integritet defineres av følgende elementer: (1) adherence: hvor godt praktikerne behersker programmets protokoll og verktøy etter opplæringen, (2) eksponering: hvor mange,

hvor lange og hvor ofte sesjonene gjennomføres, (3) implementeringskvalitet: hvilke ferdigheter, holdninger og kompetanse personalet som praktiserer programmet har, (4) responsivitet: hvor involverte og engasjerte deltakerne er og (5) program differensiering: hvor godt de ulike programkomponentene kan skilles fra hverandre (Mihalic & Irwin, 2003). Med andre ord er høy integritet et resultat av et samvirke mellom kjennetegn ved terapeut, klient og program. Det kan imidlertid være enklere å måle i hvilken grad behandlerne er lojale mot programbeskrivelsen enn hvor *kompetent* de utfører behandlingen. Mens behandlingsintegritet handler om terapeuten har gjort det han eller hun skal (eng. adherence / fidelity), handler kompetanse om hvor bra han eller hun iverksetter intervensjonen (eng. competence). Det er lite som tyder på at 100 prosent lojalitet mot behandlingsprogrammet er nødvendig for å oppnå gode resultater og mye tyder på at *kompetent implementering* innebærer at en både kan legge til og trekke noe i fra i forhold til idealmodellen uten at en forringer behandlingsresultatene. Kompetanse er praktikerens kliniske ferdighetsnivå og kan komme til uttrykk gjennom at han eller hun tar hensyn til klientens status, velger egnede tidspunkt for bruk av ulike komponenter (eng. timing) tar hensyn til konteksten for intervensjonen og generelt formidler behandlingen på en virksom måte. Og mens kompetanse forutsetter behandlingsintegritet, så impliserer ikke integritet nødvendigvis kompetanse. Waltz, Addis, Koerner & Jacobson (1993) beskriver hvordan en kan vurdere både behandlingsintegritet og kompetanse gjennom å skille mellom: (1) intervensjoner som er både unike og avgjørende for behandlingen, (b) behandlingskomponenter som er avgjørende viktig for behandlingen, men ikke nødvendigvis unike for den, (c) komponenter eller handlinger som er forenlige med behandlingen, men ikke unik eller avgjørende i forhold til den og (d) handlinger eller komponenter som frarådes eller som klart ikke er en del av terapien (dvs. at de er kontraindiserte eller tilhører en annen terapiretning).

Avveiningen mellom integritet og kompetanse henger sammen med forholdet mellom forskning og praksis. For å fastslå årsakssammenhenger mellom intervensjon og resultater i effektstudier, er det viktig å kunne dokumentere en høy grad av behandlingsintegritet slik at det er tydelig hva man tester effekten av. Fokus er da på intervensjoner som er både unike og avgjørende for at behandlingen skal gi resultater. I klinisk praksis og i effektivitets- og implementeringsstudier hvor deltakerne kan variere mer med hensyn til bakgrunnskjenntegn og tilleggsp problemer, kan terapeutenes generelle kompetanse være viktig for utfallet selv om den ikke er unik for det aktuelle behandlingsprogrammet.

Komorbiditet som utfordring for standardiserte behandlingsprogram

Standardiserte behandlingsprogram som retter seg mot enkeltdiagnoser eller avgrensede problemer har blitt kritisert for å fungere dårlig for klienter med kombinasjonsproblemer (Westen mfl., 2004). Det påpekes at sammensatte problemer og komorbiditet er regelen heller enn unntaket i ordinær praksis. De fleste som blir henvist eller søker behandling har tilleggsp problemer av forskjellig slag (APA, 2006) og problemene som i utgangspunktet var primære kan etter hvert komme i bakgrunnen mens andre kommer mer i fokus (Westen mfl., 2004). En kan heller ikke forutsette at problemer som skal behandles er additive slik at de kan behandles suksessivt med skiftende program.

Denne oppfatningen deles imidlertid ikke av alle. Andre mener at de færreste evidensbaserte programmer forutsetter at klientene bare har ett problem og at de tar høyde for klientenes tilleggsp problemer og livsforhold, selv om de også fokuserer på det henviste hovedproblemet (Weisz, Weersing & Henggeler, 2005). Dette gjelder særlig for programmer for behandling av alvorlige atferdsproblemer så som Funksjonell Familieterapi (FFT, Alexander & Parsons, 1982), Multisystemisk Terapi (MST, Henggeler mfl., 1998) og Multidimensional Treatment Foster Care (MTFC, Chamberlain, 1998). I disse programmene intervensjoner man i forhold til kognitive, så vel som familiebaserte og vennerelaterte problemer, både i familien på skolen og i nærmiljøet. Programmene er omfattende og bredspektrede, noe som er en forutsetning for at de skal virke i forhold til de multi-problemene klientene henvises for. En viktig del av å lære seg slike program består derfor i å lære seg hvordan en tilpasser behandlingen til individuelle og miljømessige kjennetegn (Weisz, Weersing & Henggeler, 2005). Påpekningen av at behandlingsprogrammer rettet mot ungdom med alvorlige atferdsproblemer har få eksklusjonskriterier ser også ut til å bli akseptert av kritikerne (Westen mfl., 2005).

Behandlingstidens lengde

Forholdet mellom korttids- og langtidsbehandling er et annet aktuelt tema i diskusjonen om evidensbaserte behandlingsprogrammer. Kritikere hevder at de fleste evidensbaserte programmene dreier seg om korttidsbehandling, og betviler at dette alltid gir de beste utbyttet for klientene (Westen mfl. 2004). De påpeker at behandlingstidens lengde ofte fastsettes av programutviklerne tidlig i prosessen og at lengden sjelden blir systematisk variert og testet. Selv om korttidsterapier har vist seg å kunne gi gode resultater, hevder Westen mfl. (2004) at tilbakefallsprosenten er stor for mange problemgrupper, enten de har fått behandling eller ikke (for eksempel depresjon). De mener at visse psykologiske symptomer vanskelig kan forstås

og behandles isolert fra personligheten og at personlighetsprosesser eller –trekk vanskelig kan endres i løpet av 6 til 16 sesjoner som er den vanlige doseringen i korttidstereapi. De viser til naturalistiske studier av psykoterapi som dokumenterer dose-response sammenhenger i langtidsbehandling, for eksempel over 1-2 år eller mer. Hvis sammenligninger mellom evidensbaserte behandlingsprogrammer og annen behandling skal være relevant hevder de at kontrollbetingelsene må være langtidsbehandling uten tidsbegrensninger.

Crits-Cristoph, Wilson og Hollon (2005) påpeker imidlertid at det er ingen ting som sier at evidensbasert behandling skal være korttidsbasert (eller manualisert), og at dette først og fremst er et empirisk spørsmål. Når det viser seg at ikke alle klienter reagerer positivt på behandlingen eller at noen får tilbakefall så har det ingenting med definisjonen, metodologien eller premissene for evidensbasert behandling gjøre. Crits-Cristoph mfl. (2005) mener at kritikerne forveksler de prototypiske studiene som har blitt publisert med premissene for evidensbasert behandling. De understreker at det er ingen ting i veien for at evidensbasert behandling kan være langtidsbasert, integrerende eller eklektisk. Korttidsbehandling er først og fremst utviklet av klinikere og forskere som har erfart at slik behandling kan hjelpe mange klienter. Det er dessuten et fornuftig utgangspunkt fordi mange klienter ikke blir i behandling særlig lenge. Etter å ha gjennomgått en database med 6000 behandlinger av voksne pasienter fant Weisz mfl. (2005) at de fleste terapiene var av kort varighet med mindre enn fem sesjoner i gjennomsnitt (Hansen, Lambert & Forman, 2002). Samtidig understreker de at hvis klienter ikke har utbytte av korttidsintervensjoner, så er det ingen ting i veien for å drive langtidsbehandling. Men hvis valget står mellom langtidsbehandling av ukjent kvalitet for alle klienter eller korttidsbehandling med kjent effekt for dem som har utbytte av det, så kan den sistnevnte strategien ha mye for seg (Crits-Cristoph mfl., 2005). Det bør også nevnes at behandlingsavslutningen innenfor flere evidensbaserte behandlingsprogrammer ikke er tidsavgrenset, men kriteriebasert (for eksempel MST og PMTO). Det vil si at behandlingen avsluttes når behandlingsmålene er nådd eller at klienten fungerer bedre sosialt. Den norske PMTO studiene som presenteres i denne avhandlingen viste for øvrig at doseringen ikke korrelerte med resultatvariablene og at det derfor ikke var noen sammenheng mellom behandlingstidens lengde og resultatene for klientene (Ogden & Amlund-Hagen, 2008).

Som det framgår av de foregående avsnittene kan mange av de omstridte temaene innenfor evidensbasert behandling tilbakeføres spørsmålet om hva det skal forskes på. Motsetningen består mellom behandling som tar utgangspunkt i standardiserte forskningsbaserte programmer og behandling som i større grad vektlegger det som er felles for flere programmer, modeller eller metoder. Debatten om evidensbasert behandling har

særlig kretset rundt spørsmålet om hvor viktig spesifikke programfaktorer er for behandlingsutfallet sammenlignet med terapeutiske fellesfaktorer som terapeutens kjennetegn og kompetanse, klientenes kjennetegn og den terapeutiske alliansen (Sprenkle & Blow, 2004; Wampold, 2001).

Programfaktorer og terapeutiske fellesfaktorer

Representanter for den evidensbaserte tilnærmingen mener at effektforskningen viser at det finnes meningsfylte forskjeller mellom spesifikke behandlingsprogrammer for spesifikke problemer (Chambless & Ollendick, 2001; Sexton mfl., 2004). Andre forskere har imidlertid tatt til orde for at det ikke er kvantitative forskjeller mellom ulike tilnærminger og at uspesifikke eller fellesfaktorer forklarer det meste av variansen i behandlingsresultater (Lambert, 1992; Shadish mfl., 1993; Wampold mfl., 1997; Wampold, 2001). Fellesfaktorene tilhører ikke noen spesiell teoretisk tilnærming eller behandlingsmodell, men står i motsetning til spesifikke faktorer som er unike for et bestemt program eller for en teoretisk tilnærming eller modell (Waltz mfl. 1993). Fellesfaktor tilnærmingen har en lang historie og ble blant annet beskrevet av Rosenzweig i 1936 og av Jerome Frank i 1961. Fellesfaktor perspektivet oppsto i den individuelle psykoterapi litteraturen fordi det i motsetning til de tradisjonelle terapiskolene kunne forklare motsetninger mellom teori og praksis, og ble også støttet gjennom flere meta-analyser (Sexton mfl., 2004). Meta-analyser viste at individualterapi virket sammenlignet med ingen behandling, men at det ikke var mulig å identifisere forskjeller i behandlingseffekt mellom brede teoretiske tilnærminger (Smith, Glass og Miller, 1980; Shadish & Baldwin, 2002; Wampold mfl., 1997).

Innenfor fellesfaktor tilnærmingen kan det skilles mellom et radikalt eller smalt perspektiv og et moderat eller bredt perspektiv på terapeutiske fellesfaktorer.

Det radikale perspektivet på fellesfaktorer

Det radikale eller *smale* perspektivet beskriver fellesfaktorene som ikke-spesifikke aspekter ved behandlingsmodeller så som å skape meningsendringer eller 'reframing' (Lambert, 1992). I en meta-analyse fikk Wampold mfl. (1997) bekreftet hypotesen om at det ikke var noen forskjeller i behandlingseffekt mellom ulike former for psykoterapi. Resultatet ble fortolket som at det var de uspesifikke fellesfaktorene som kvaliteten på terapeut-klient relasjonen som forklarte mesteparten av variansen i behandlingsresultater. Wampold (2001) utformet en *kontekstuell modell* for psykoterapi med utgangspunkt i Frank og Franks (1991) beskrivelse

av felleselementer i psykoterapi: et emosjonelt ladet tillitsforhold til en hjelpende person, en terapeutisk setting hvor klienten tror at den profesjonelle er i stand til å gi hjelp, en terapeut som formidler et troverdig rasjonale og en plausibel teori for å forstå pasientens symptomer og at terapeuten tilbyr et troverdig rituale eller prosedyre for å håndtere symptomene. På grunnlag av denne beskrivelsen definerte Wampold mfl. (1997) fire fellesfaktorer som kunne forklare det meste av variansen i kliniske resultater; den terapeutiske relasjonen som forklarer cirka 30%, placebo effekten som forventninger og håp om endring forklarer cirka 15%, modell eller terapiteknikk som er utviklet for å løse et problem forklarer cirka 15% og faktorer utenfor terapien forklarer cirka 40% av variasjonen i behandlingsresultatene. Disse prosentene bygger på et estimat i en kvalitativ analyse av psykoterapi forskning der Lambert (1992) foreslo en firefaktors modell basert på felles endringselementer på tvers av teorier. Selv om prosentandelene var estimater, hadde de i følge Hubble mfl. (1999) blitt sitert i litteraturen som om det dreide seg om variablenes faktiske andel forklart varians (Duncan & Miller, 2005). Wampold mfl. (1997) redegjorde heller ikke for hvordan disse beregningene ble utført og empiriske studier viste andre tall (Horvath & Symmonds, 1991).

I følge Wampold (2001) forklarte programfaktoren bare en liten del av variasjonen i behandlingsutfall, og det var først og fremst terapeutkvalitetene som var avgjørende for en vellykket behandling. Men meta-analyser er som andre analytiske metoder avhengig av dataene de bygger på og Wampolds analyse tok utgangspunkt i brede og generelle terapeutiske skoleretninger eller globale komponenter som kognitive eller atferdsmessige komponenter (Sexton mfl., 2004). Verken disse eller andre tidligere meta-analyser kodet for intervensjonstyper som fanget opp nyere spesifikke intervensjonsmodellene for spesifikke problemer som MST, FFT og MTFC (Sexton & Alexander, 2002). I lys av disse vanskene kan ikke konklusjonen om manglende forskjeller mellom kliniske tilnærminger bekreftes av Wampolds (2001) eller tidligere meta-analyser (Sexton mfl., 2004).

Chambless og Ollendick (2001) kritiserte også Wampold og kollegers (1997) meta-analyse på tre punkter. For det første, fokuserte meta-analysen i hovedsak på et begrenset utvalg klienter, blant annet college studenter og polikliniske pasienter med problemer med angst, depresjon, spiseforstyrrelser og lignende. Sammenligningen omfattet ikke behandling av kronisk psykisk syke, klienter med utviklingshemning eller atferdsforstyrrelser. For det andre manglet meta-analysen spesifisitet og inneholdt lite forskning om barn og unge. Viktige forskningstidsskrifter manglet også i oversikten. Og for det tredje har meta-analyser av behandlingsforskning med barn og ungdom kommet til andre konklusjoner om forholdet mellom programfaktorer og fellesfaktorer (Chambless & Ollendick, 2001). Weisz mfl. (1995)

fant for eksempel at atferdsorientert behandling (inkludert parent management training) var mer effektiv enn ikke atferdsorientert behandling (jf også Casey & Berman, 1985).

Wampold (2001) illustrerte sine funn med et utsagn fra Alice i eventyrland kalt "the dodo-bird verdict" som formidler at "alle har vunnet og alle skal ha premie". Metaforen hadde tidligere blitt brukt av Rosenzweig (1936) og Luborsky mfl. (1975) for å beskrive at det er minimale resultat-forskjellene mellom ulike terapimodeller. Weisz mfl. (2005) drøftet også funnene i en meta-analyse i forhold til "the dodo bird verdict". På den ene siden tilbakeviste de påstanden fordi resultatene deres ganske konsekvent viste at evidensbaserte behandlingsprogram med definerte komponenter ga bedre resultater enn ordinær behandling. Men de hevdet også at siden den samlede eller gjennomsnittlige effektstørrelsen i meta-analysen var moderat så var funnene ikke helt uforenlig med "the dodo bird notion". Sprenkle & Blow (2004) har også bidratt til denne debatten og konkluderte med at 'dodo bird' utsagnet var uheldig valgt, selv om både de selv og andre talsmenn for fellesfaktor tilnærmingen hadde brukt det tidligere (Sprenkle, Blow & Dickey, 1999). Deres begrunnelse var at hvis en konkluderer med at ingen behandling er bedre enn noen annen, så vil det kunne gi inntrykk av at det ikke betyr noe hva man gjør i terapi, og at en ikke kan skille mellom empirisk validerte terapier og 'tarot kort' eller 'ouija boards' (Sprenkle & Blow, 2004).

Den moderate fellesfaktor tilnærmingen

Den moderate fellesfaktor tilnærmingen har et bredere perspektiv enn det foregående og fellesfaktorene omfatter klienten, terapeuten, behandlingsrelasjonen, forventninger og andre behandlingsvariabler som ikke er spesifikke for en bestemt modell (Sprenkle & Blow, 2004). Men for noen definerte og krevende problemer kan også spesifikke programfaktorer være utslagsgivende slik at bestemte behandlingsprogrammer (eller –modeller) kan være bedre egnet for visse problemer, klienter, forhold og terapeuter enn andre. Fellesfaktorer og spesifikke faktorer kan i følge dette perspektivet best beskrives som et både-og, og ikke som et enten-eller. Forfatterne hevder at det finnes resultatforskjeller mellom virksomme psykoterapier, men fellesfaktorer og felles endringsmekanismer som ligger bak de fleste vellykkede behandlingsformer forklarer langt mer av variansen i resultatene enn de unike aspekter ved behandlingsmodellene. I følge dette moderate fellesfaktor perspektivet har fellesfaktorene ikke har fått den plassen de fortjener og blir ofte oversett i klinisk forskning (Sprenkle & Blow, 2004).

Terapeutiske fellesfaktorer

Behandlingsforskningen underbygger at kjennetegn ved behandlingsmetoden (Nathan & Goreman, 2007), den enkelte behandler (Wampold, 2001), behandlingsrelasjonen (Norcross, 2002) og klienten (Bohart & Tallman, 1999) bidrar til positivt behandlingsutfall for klienter (APA, 2006). Fordi terapi er en interaksjonsprosess, så kan det være vanskelig å skille ut effekten av enkeltvariabler. Likevel har den kliniske forskningslitteraturen beskjeftiget seg mye med den relative betydningen av programmer, terapeuter og klienters kjennetegn, inkludert den terapeutisk alliansen, for klientenes utbytte (APA, 2006). Blant de viktigste fellesfaktorene inngår terapeutfaktoren, klientfaktoren og deres felles bidrag til den terapeutiske alliansen.

Betydningen av terapeuten for behandlingsutfall

Gode terapeuter er gode terapeuter, uavhengig av hvor de tilegner seg ferdigheter fra (Sprenkle & Blow, 2004). Gode terapeuter skaper en sterk allianse med klientene, de tilbyr et aktivitetsnivå som samsvarer med klientenes forventninger og preferanser, de er kreative i å finne nye måter å nærme seg problemer på, de foreslår nye måter å lære tilpasningsferdigheter på, og de er personlig godt integrerte (Beutler mfl., 1998; 2004). Men utover terapeutenes bidrag til behandlingsalliansen, har det vært forsket lite på hva som skiller kompetente terapeuter fra de mindre gode. Det finnes ikke klare regler for hvordan en skal velge ut terapeuter til å delta i kliniske forsøk, og terapeutfaktorens betydning blir sjelden rapportert. Det kan blant annet forklares med at økt interesse for evidensbaserte behandlingsmetoder har resultert i mindre oppmerksomhet rundt terapeutfaktorer. I kliniske effektstudier har terapeutfaktoren ofte blitt behandlet som feilkilde heller enn som varianskilde og man har forsøkt å kontrollere for variasjoner i terapeutfaktorer ved bruk av behandlingsprotokoller og måling av behandlingsintegritet. Det er imidlertid betydelig evidens for terapeutforskjeller, selv i prosjekter der en tester standardiserte behandlingsprogrammer. For å fortolke resultatene fra behandlingseffekt studier kan det være nyttig å vite hvordan en velger ut terapeutene, hvor mange de er, hvem de er og deres teoretiske orientering. Videre kan det være av interesse å vite hvor kompetente, engasjerte og forpliktet de er i forhold til behandlingsmetoden, hvor stor behandlingsmengde de har, hvor ofte de ser klientene, samt hvor mye opplæring og veiledning de får i metoden de bruker. Manglende resultater som skyldes at en intervensjon ble dårlig implementert kan blant annet skyldes at terapeutene fikk mangelfull eller utilstrekkelig opplæring og veiledning (Chambless & Hollon, 1998). Det kan

lett forekomme når terapeutene er ukjent med behandlingsopplegget eller har preferanse for en annen tilnærming enn den de skal bruke.

Resultater som viser at terapeutenes erfaring betyr lite for resultatene, kan ha sammenheng med at erfaring blir målt som kvalifikasjoner (eng. 'credentials') heller enn kompetanse, eller målt som antall års erfaring med en hvilken som helst psykologisk behandling, heller enn opplæring og erfaring med en bestemt terapiform (Chambless & Hollon, 1998). Mye tyder på at terapeuteffektene er mindre i kontrollerte kliniske forsøk sammenlignet med naturalistiske studier av klinisk praksis, fordi terapeutene får mer opplæring og veiledning enn terapeuter i ordinær praksis, men også fordi det ofte dreier seg om utvalgte dyktige terapeuter (Crits-Christoph mfl., 1991). Terapeuter som deltar i kliniske forsøk kan også ha hatt fordelene av å arbeide over lengre tid med et avgrenset problem på en definert måte. Slik konsentrert innsats kan gi bedre resultater enn ordinær praksis. Derfor kan det være viktig å vite om terapeutene som jobber med klienter i kontrollgruppen har like gode rammer og støtte i sitt arbeid, om de er like motiverte og engasjerte, og om de ser klientene like ofte som terapeutene som behandler eksperimentgruppen (Westen mfl, 2004; 2005).

Betydningen av klientkjennetegn for behandlingsutfall.

Ifølge Duncan og Miller (2000) er klienten den 'egentlige helten' i psykoterapi og de understreker at terapi bare virker i den grad den aktiverer klientens naturlige forbedringspotensiale. Tallman og Bohart (1999) mener at grunnen til at flere behandlingsmodeller kan fungere like bra er klientenes evne til å bruke det de får tilbud om og at dette er viktigere enn forskjellene i teknikker og tilnærminger. Lambert (1992) hevder også at klientforskjeller og ekstraterapeutiske faktorer er de viktigste varianskilden i psykoterapi. *Klientfaktorer* er kjennetegn ved klienten så som motivasjon og forpliktelse i forhold til endring, indre styrke og (religiøs) tro mens *terapeutiske fellesfaktorer* er ingredienser i klientens liv og miljø som påvirker endring, så som sosial støtte, engasjement i lokalmiljøet og stressende hendelser. Klienter må blant annet være villige til å akseptere behandlingstilbudet og klare å opprettholde sitt engasjement under behandlingen. Problemer oppstår når klienter faller fra, eller deltar i varierende grad. Det er grunn til å tro at de som fullfører behandlingen skiller seg ut i forhold til de som faller ut, og i frafallsgruppen kan det være forskjeller mellom de som faller ut tidlig og de som slutter på et senere tidspunkt i terapien. Så selv om behandlerne anstrenger seg for å gi det samme tilbudet til alle vil ulike klientgrupper ha mottatt ulike behandlingsdoser (Lipsey, 2000). Klientkjennetegn som funksjonsnivå ved inntak (Elkin mfl., 1995) og evnen til å inngå i en allianse med terapeuten

(Krupnick mfl., 1996) og har også vist seg å være viktige prediktorer for behandlingsutfall (Chambless & Ollendick, 2001).

Betydningen av terapeutisk allianse for behandlingsutfall

Både terapeuten og klienten bidrar til å skape den terapeutiske relasjonen når de sammen fokuserer på behandlingen. Terapeutisk allianse blir ofte framhevet som en viktig fellesfaktor og endringsmekanisme og har vist seg å kunne forklare opptil 30% av resultatvariansen i psykoterapi (Kazdin, 2008). Blant evidensbaserte programmer i par- og familierapi som har hatt mest empirisk støtte som Funksjonell familierapi (FFT) og Multisystemisk Terapi (MST) legges det vekt på å bygge en sterk terapeutisk allianse (Sprenkle & Blow, 2004). Bordin (1979) beskriver alliansen som bestående av tre elementer: (1) *bånd*, som er den affektive kvaliteten av klient-terapeut relasjonen og omfatter dimensjonene tillit, omsorg og involvering, (2) *oppgaver*, som handler om i hvilken grad klienten og terapeuten er bekvemme med hovedaktivitetene i terapien og om klienten finner dem troverdige, og (3) *mål*, som handler om i hvilken grad terapeut og klient arbeider mot forenlig mål. Forskning om allianse har omfattet alle disse tre variablene.

Shirk og Karver (2003) fant at den terapeutiske relasjonen var en rimelig robust og konsistent forklaringsfaktor i en meta-analyse av 23 studier (hvorav 18 publiserte) av sammenhengene mellom den terapeutiske relasjonen og kliniske resultater for barn og ungdom. Den terapeutiske relasjonen var moderat korrelert med resultatene, både på tvers av behandlingsformer og på ulike alders- og utviklingstrinn. De fant at det var vanskeligere, men også mer avgjørende å etablere en terapeutisk relasjon i arbeidet med barn med eksternaliserende enn med internaliserende problemer. I studier av "Parent Management Training" (PMT) der terapeutens har begrenset direkte kontakt med barnet, er relasjonen mellom terapeut og foreldre sentral. Det viste seg derfor at sammenhengen mellom relasjon og resultater var sterkere for foreldrebehandling enn for behandling av barn, men det gjorde ingen forskjell om behandlingen var atferdsorientert eller ikke. Den terapeutiske relasjonen forklarte heller ikke mer varians i ikke-manualiserte enn i manualiserte behandlinger, og den var heller ikke mer utslagsgivende i ordinær praksis enn i forskningsbasert praksis (selv om det var tendenser i den retningen) (Shirk & Karver, 2003).

Terapeutisk allianse bør i følge Shirk og Karver (2003) måles før avslutningen av terapiene og helst flere ganger underveis, for å unngå at svarene påvirkes av at foreldrene registrerer om terapien går i positiv eller negativ retning. Mange av studiene i deres meta-analyse målte imidlertid den terapeutiske relasjonen ved slutten av terapien. En annen

metodisk svakhet ved studiene var at sammenhengen kunne være overvurdert fordi foreldrene rapporterte både om den terapeutiske relasjonen og barnas atferd.

Terapeutisk allianse blir ofte framhevet som en viktig fellesfaktor som kan bidra til å forebygge frafall i behandlingen av barn (Armbruster & Kazdin, 1994) og den kan forklare en betydelig andel av variansen i behandlingsutfall. Men den forklarer ikke hvilke terapeutiske endringsmekanismen som ligger bak (Kazdin, 2008; Sexton mfl., 2004). Alliansen forklarer ikke hvorfor behandlingen leder til endring, eller hvor viktig den er som påvirkningsfaktor i terapi. Terapeutisk allianse kan selv bli forklart av én eller flere andre variabler som for eksempel symptomendring som forekom før eller samtidig som alliansen begynte å ta form. Den kan også forklares av felles metodevarians eller av kjennetegn ved klienten før de startet i behandling. Så selv om terapeutisk allianse kan være en viktig kilde til resultatvariasjon, så kan en overfortolke eller feiltolke dens betydning når en bare fokusere på mengden av forklart varians (Kazdin, 2008). Oppsummert tyder forskningen på at den terapeutiske relasjonen i behandlingen av alvorlige atferdsproblemer hos barn og unge har en uklar status, og at mer og bedre designet forskning er nødvendig for å kunne estimere den terapeutiske relasjonens betydning for de kliniske resultatene.

En prosessmodell for terapeutisk endring

Sexton mfl. (2004) hevder det er mye som tyder på at fellesfaktorene har noe for seg og at det kan finnes noe felles i all variasjonen: "Det er sentrale og felles faktorer som bidrar til vellykkede resultater på tvers av tilsynelatende forskjellige teori og praksismodeller" (s. 134). Dessuten peker forskningsresultatene mot felles heller enn spesifikke sett av aktive ingredienser både i individualterapi og familierapi. Den sterkeste støtten til fellesfaktor tilnærmingen kommer fra meta-analyser som har konkludert med at det ikke er forskjeller mellom terapeutisk intervensjoner. Men svakheten ved disse meta-analysene er at de sammenlignet brede teoretiske behandlingsmodeller og komponenter heller enn spesifikke programmer for spesifikke problemer.

Sexton mfl. (2004) er opptatt av hvilke endringsprosesser som gjør terapi virksom. Heller enn å identifisere faktorer som er forbundet med endring vil de forklare prosesser og mekanismer for endring. Fellesfaktorer forklarer med andre ord ikke kompleksiteten i endringsprosessene og som endring finner sted gjennom. De er verken operasjonelt definerte, kontekstualisert innnefor rammene av den kliniske prosessen eller forklart med hensyn til hvordan de fremmer resultater. Mange av fellesfaktorene er resultatene av en udefinert relasjonell og interpersonlig prosess heller enn de terapeutiske endringsmekanismene de

impliserer. Et eksempel er ”håp og forventninger” som antagelig bidrar til forandring på en eller annen måte, men de beskriver ikke den underliggende mekanismen for terapeutisk forandring, de bestemte terapeut handlinger som aktiverer dem, hvordan de inngår i prosessen, eller når en kompleks endringsprosess om håp og forventninger er viktigst. Felles endringsprosesser er derfor et alternativ til fellesfaktorer og hvis forskere skal forbedre endringspotensialet i psykoterapi må de rette oppmerksomheten mot viktige aspekter ved endring som å intensivere klientenes motivasjon, øke deres involvering i terapien, øke deres tro på at endring kan finne sted og fjerne det klienten oppfatter som hindringer for forandring (Hanna, 1996). Sexton, Alexander mfl. (2004) hevder at felles endringsmekanismer i terapier ikke er det samme som fellesfaktorer, men dreier seg om endringselementer som ikke kan beskrives uavhengig av modellen de er virksomme innenfor. Evidensbaserte behandlingsmodeller som integrerer praksis, forskning og teori på måter som forbedrer praksis har forskjellige teoretiske prinsipper og terapeutiske prosesser, men likevel noen fellestrekk. De er velformulerte systematiske behandlinger hvor: (1) en behandler klinisk meningsfylte problemer, (2) et sammenhengende begrepssystem danner grunnlag for de kliniske intervensjonene, (3) spesifikke intervensjoner blir beskrevet detaljert med en presisering av hvilke terapeutkvaliteter som er nødvendige for å kunne følge dem, (4) prosessforskning identifiserer hvordan endringsmekanismene fungerer og (5) forskning om resultater viser hvor bra de fungerer (Alexander, Sexton & Robbins, 2002; Sexton mfl., 2004).

Fellesfaktorer og endringsmekanismer

Oppsummert hevder Sexton mfl. (2004) at fellesfaktor tilnærmingen overforenkler de komplekse prosessene i familierapi og har ført til en polarisering hvor en enten velger fellesfaktorer eller en spesifikk klinisk modell. De foreslår at en betrakter fellesfaktorer som nødvendige, men ikke tilstrekkelige ingredienser i spesifikke og omfattende modeller for terapeutisk endring.

De terapeutiske fellesfaktorene kan stilles opp som motsetning til programspesifikke faktorer, men de kan også integreres i randomiserte, kontrollerte design. Vanligere er det at fellesfaktorer studeres i naturalistiske ’process-outcome’ studier der prediksjonsstyrken i ulike variabler analyseres og sammenlignes. I kliniske studier hvor behandlingsprogrammer sammenlignes med sine kontrollbetingelser kan imidlertid fellesfaktorene innarbeides i designet. Både terapeut og klientkjennetegn kan inngå som modererende variabler, og avhengig av på hvilket tidspunkt i behandlingen den måles på, kan terapeutisk allianse være både moderator og mediator. Utviklingen innenfor det randomiserte designet har også vist at

fellesfaktorer i økende grad inngår blant undersøkelsesvariablene (LaGreca, Silverman & Lochmann, 2009). En slik utvikling reduserer nødvendigvis ikke meningsforskjellene om den relative betydningen av programfaktorer og fellesfaktorer, selv om det radikale perspektivet på fellesfaktorer står mer i kontrast til den evidensbaserte tilnærmingen enn det moderate perspektivet gjør.

Er programbasert behandling mer eller mindre effektiv enn behandling basert på klinisk skjønn?

Når behandling basert på evidensbaserte programmer sammenlignes med behandling basert på klinisk skjønn hevder kritikere at terapeuter opplært i standardisert behandling ikke er fleksible nok. De klarer ikke å avvike fra behandlingsmanualene når det er nødvendig i en bestemt sak og dermed kan høy grad av behandlingsintegritet gi dårligere resultater enn individuelt tilpasset behandling (Chambless & Ollendick, 2001). Persons (1991) hevdet for eksempel at kontrollerte kliniske forsøk hindrer klinikere i å bruke sine vurderinger i arbeidet med å skreddersy behandlingen til klientenes spesifikke behov og at resultatene ville ha blitt bedre hvis praktiseren kunne drive en fleksibel behandling.

For å teste forholdet mellom standardisert, programbasert behandling og behandling basert på klinisk skjønn bør en undersøke hvor godt terapeuter opplært i et standardisert program lykkes i ustrukturert klinisk praksis (Chambless & Ollendick, 2001). Det foreligger enkelte randomiserte studier der en kontrollerte for terapeutfaktoren, det vil si der en analyserte hvilke resultater terapeuter opplært i evidensbaserte behandlingsprogram oppnådde i saker der de kun anvendte klinisk skjønn (Jacobson mfl., 1989; Schulte, 1992; Emmelkamp, Bouman & Blaauw, 1994). For å unngå at terapeutfaktoren skulle bidra til resultatene, administrerte de samme terapeutene behandlingen i henholdsvis en standardisert og en individualisert behandlingsbetingelse. Klientene ble tilfeldig fordelt på behandlingsbetingelsene og ikke i noen av studiene var fleksibel behandling statistisk overlegen i forhold til standardisert behandling. Terapeutene oppnådde bedre resultater når de fulgte en fast protokoll enn når de forsøkte å individualisere behandlingen. Studiene omfattet imidlertid parterapi og behandling av fobier hos voksne, så spørsmålet om resultatene var gyldige for barn og unge er en utfordring for framtidig forskning (Chambless & Ollendick, 2001).

Behandling av barn skiller seg fra behandling av voksne på måter som underbygger at praksis i forskningsstudier av barn og ungdom har større overføringsverdi til ordinær praksis

enn studier av voksne. Det er som regel foreldre eller andre voksne heller enn barna selv som tar initiativet til behandlingen, og henvisningen er ofte utløst av et uvanlig høyt problemnivå hos barna kombinert med mange tilleggsproblemer (komorbiditet) samt et høyt stressnivå hos foreldrene og i familien. Et slikt utgangspunkt kan føre til upredikerbare hendelser eller uventede endringer hos klient eller i situasjonen fra den ene terapitimen til den neste. Faren for at klientene ikke møter som avtalt eller faller ut av behandlingen er betydelig, fordi både foreldre og barn skal involveres (Weisz mfl., 2005). Det er derfor behov for en sterk arbeidsallianse for å opprettholde frammøte og oppnå gode resultater. Usikkerheten og upredikerbarheten i behandlingen av barn og unge synes å gjelde enten det dreier seg om effektstudier eller effektivitetsstudier.

Weisz, Jensen-Doss og Hawley (2006) identifiserte 32 effektivitetsstudier med barn og ungdom der empirisk støttede terapier ble sammenlignet med andre aktive behandlingstilbud i ordinær praksis. Meta-analysen viste at den gjennomsnittlige effektstørrelsen for evidensbasert behandling i forhold til aktive intervensjoner i vanlig praksis var $d = .30$. Hensikten var å finne undersøkelser med kontrollbetingelser der terapeutene kunne anvende sine kliniske vurderinger i en for dem riktig behandling, og uten å være bundet av manualer. Deltakerne i disse studiene ble randomiserte til evidensbasert behandling eller vanlig behandling ('usual care' som besto av psykoterapi, rådgivning eller saksbehandling som en del av det ordinære tjenestetilbudet til barn og unge). Dette resulterte i en meta-analyse av 32 studier (23 publiserte artikler og 9 avhandlinger) som var gjennomført mellom 1965 og 2004 med barn i aldergruppen 3-18 år. Av disse fokuserte 17 på behandling av kriminalitet/ alvorlige normbrudd og rusmisbruk, 12 på alvorlige atferdsproblemer og 3 på internaliserte problemer (programmet som oftest ble testet i denne meta-analysen var Multisystemisk Terapi med 6 studier). Meta-analysene av behandlingsstudier basert på klinisk skjønn hadde en gjennomsnittlig effektstørrelse rundt null (for eksempel Weisz, 2004; Weisz mfl. 1995).

En av disse studiene var en evaluering av behandling av problemungdom i en lokal klinikk (Weiss mfl., 1999). Terapeutene kunne fritt velge hvilken metode de ville bruke, og behandlingene besto i gjennomsnitt av 60 individuelle sesjoner, inkludert 18 foreldresesjoner og 13 skolekonsultasjoner. Behandlingen som ble gitt etter terapeutenes eget metodevalg viste at resultatene for ungdommene var de samme som for ungdommer som ble tilfeldig fordelt på en kontrollgruppe som bare fikk skolefaglig hjelp. Andre undersøkelser der en kombinerte ulike former for ordinær behandling (TAU – treatment as usual) i en tiltakskjede (eng. systems of care) ga heller ikke bedre resultater (Bickman, 1996; Bickman mfl., 1995; Bickman, Noser & Summerfelt, 1999).

Weisz mfl. (2006) konkluderte sin meta-analyse med at evidensbasert behandling ga bedre resultater enn etablerte (ordinære) tjenester i direkte randomiserte sammenligninger. Meta-analysen viste at uansett hvordan en sammenlignet, så kom evidensbasert behandling bedre ut enn ordinære tjenester. Men ikke uventet var innholdet i de ordinære behandlingstiltakene ofte så mangelfullt beskrevet at det var vanskelig å klassifisere dem på en adekvat måte. Noen av de evidensbaserte programmene hadde de høyeste effektstørrelsene i meta-analysen, blant annet MST, PMT, Problem Solving Skills Treatment og Cognitive Self-instruction Training. Men ikke alle resultatene gikk i favør av evidensbasert behandling. Noen former for ordinær behandling som ble gjennomført av dyktige klinikere syntes å ha gunstige effekter (selv om ikke alle forskjellene var signifikante). Evidensbasert behandling kan med andre ord også praktiseres så dårlig at den ikke gir forventet resultat. Blant forbedringsmulighetene nevnte forfatterne bedre håndtering av tilleggsproblemer eller komorbiditet, større fleksibilitet i terapeutenes bruk av evidensbaserte komponenter, bedre samordning av evidensbasert behandling med annen, ordinær behandling og intervensjonsforskning der programutviklere samarbeid tett med praktikere (Weisz mfl., 2006).

Kazdin (2008) oppsummerte at klinisk skjønn i form av kliniske vurderinger, beslutninger og ekspertise ikke har vært særlig nyttige for å integrere behandlingsrelevant informasjon. Klinisk skjønn skal bidra til å skreddersy behandlingen til den enkelte klients behov. Men han fant lite empirisk støtte for at slike kliniske vurderinger var reliable, dvs. konsistente hos den samme terapeut over tid eller konsistente mellom ulike terapeuter på samme tidspunkt. Beslutninger basert på klinisk skjønn har etter Kazdins (2008) vurdering ikke gitt mer positive behandlingsresultater enn om en benyttet mindre fleksible beslutningsmodeller. Presentert for den samme klientsaken, ville ulike terapeuter velge ulike behandlingsplaner. Generalisering av resultater i klinisk praksis er også et problem; hvis hver klient er unik, vil det ved begynnelsen av behandlingen være vanskelig å generalisere fra tidligere klienter. Og hvis man virkelig mener at hver klient er så unik at det påvirker beslutninger i behandlingen, så er det vanskelig å vite hvordan klinikere skal kunne ta velbegrunnede beslutninger (Kazdin, 2008).

Oppsummert viste sammenligningene mellom programbasert behandling og behandling basert på skjønn at standardisert behandling kom bedre ut når en holdt terapeutfaktoren konstant. Evidensbasert behandling kom også best ut i meta-analyser der en sammenlignet empirisk støttede terapier med andre aktive behandlingstiltak for barn og unge. Imidlertid påpekte

forfatterne av meta-analysen at resultatene var moderate og ikke konsekvent i favør av evidensbasert behandling (Weisz mfl., 2006). I kombinasjon med metodekritikken som ble rettet mot meta-analysen, kan det tyde på at det er behov for ytterligere forskning der en sammenligner effekten av programbasert behandling med behandling basert på skjønn. Kazdins (2008) konklusjoner tyder også på at evidensbasert behandling er skjønnsbasert behandling overlegen, og han peker også på at det knytter seg reliabilitets- og generaliseringsproblemer til behandling basert på klinisk skjønn.

Oppsummerende diskusjon

Evidensbasert praksis i form av forhåndsstrukturerte program har vært kontroversielle i praksis- og forskningsmiljøer, og det er særlig kritikerne som har dominert debatten både i Norge og internasjonalt (Angell, 2003; Bergmark & Lundström, 2007; Duncan & Miller, 2005; Ekeland, 1999; Haugsgjerd, Jensen, Karlsson & Løkke, 2009; Westen mfl. 2004, 2005). Som det framgår av dette kapitlet kan kritikken av evidensbaserte intervensjoner samles rundt noen hovedtemaer, men innholdet varierer også noe avhengig av om den framsettes i forhold til psykoterapi, sosialt arbeid eller andre former for intervensjon. I psykoterapifeltet har kritikerne blant annet vært opptatt av at evidensbaserte behandlingsprogrammer er psykologiens forsøk på å etablere et alternativ til farmakologisk behandling gjennom korttidsbaserte terapier med et definert innhold behandler definerte problemtilstander. I sosialt arbeid har kritikken i større grad rettet seg mot at evidensbasert behandling er et uttrykk for et kognitivt og atferdsorientert hegemoni, fordi disse behandlingsformene i større grad enn andre kan tilpasses rammene for forhånds-strukturert korttidsterapi. Blant de vanligste innvendingene mot evidensbasert praksis i sosialt arbeid er argumentet om at den ignorerer klinisk ekspertise og klientenes verdier og preferanser. Til dette svarer Gibbs og Gambrill (2002) at evidensbasert praksis handler om å integrere klinisk ekspertise med forskningsfunn og informasjon om klientenes kjennetegn og situasjon og de advarer mot prematurt å avvise evidensbasert praksis i sosialt arbeid (Gibbs & Gambrill, 2002).

I psykoterapiforskningen har mye av uenigheten dreid seg om den relative betydningen av spesifikke programeffekter og terapeutiske fellesfaktorer. Representanter for fellesfaktor tilnærmingen kan deles i en radikal gruppe som avviser at det finnes programeffekter (Lambert, 1992; Luborsky mfl., 1975; Wampold mfl, 2007; Wampold, 2001) og én som anerkjenner at det finnes programeffekter selv om fellesfaktorenes betydning etter deres

oppfatning ofte undervurderes (Sprenkle & Blow, 2004). I tillegg finnes det forskere som hevder at selv om fellesfaktor perspektivet er intuitivt tiltalende, så representerer det et overforenklet perspektiv på de komplekse prosessene som forekommer i behandling (Sexton mfl., 2004).

Det kan imidlertid virke kunstig å stille metode og terapeut opp mot hverandre slik det radikale perspektivet gjør. Men perspektivet understreker imidlertid viktigheten av å rekruttere gode terapeuter, eller av å gi dem gode utviklingsmuligheter, blant annet gjennom utdanning og veiledning. I praksis er det ikke et spørsmål om enten eller, men snarere om et "både-og". Selv om det er noe sannhet i utsagnet om at "alle terapiformer kan få gode effekter med rett terapeut" (Nathan & Gorman, 2007), så betyr ikke det at alle terapiformer er like virksomme. En dyktig terapeut med et godt program oppnår rimeligvis bedre resultater enn en dyktig terapeut med et dårlig program. Dessuten er det mye som taler for at en dyktig terapeut vil være mer enn gjennomsnittlig interessert i å sette seg inn i behandlingsprogram som både er konkret beskrevet og som gjennom kontrollert evalueringsstudier har vist seg å gi gode og varige resultater. Det er neppe noen som betviler at terapeutens kjennetegn har innflytelse på hvor godt evidensbaserte behandlinger virker eller hvor gode resultater de gir. Men det må også tas hensyn til at den terapeutiske relasjonen kan påvirkes av at terapeuter opplever at de kan anvende behandlingstiltak som er relevante, motiverende og som kanskje de selv og andre har hatt positive erfaringer med. Et virksomt og troverdig program kan også være en viktig rekrutteringsfaktor i miljøer som arbeider med et definert problemområde som for eksempel alvorlige atferdsproblemer.

Forskningen om evidensbaserte behandling har lenge vært opptatt av programeffekter som hovedeffekter, men i har også rettet oppmerksomheten mot forhold som modererer behandlingsresultater og mot endringsmekanismer (Cook, 2003; LaGreca, Silverman & Lochman, 2009). Randomiserte kontrollerte undersøkelser kan inneholde subgruppeanalyser som viser at at kjennetegn ved barn (for eksempel alder, kjønn og problembelastning) og familier (for eksempel familiestruktur, inntekt og utdanning) modererer behandlingsresultatene (Ogden & Amlund-Hagen, 2008). Det samme gjelder for terapeutkjennetegn, behandlingsintegritet og behandlingsallianse. I behandlingen av barn og unge med atferdsproblemer vil kjennetegn ved barnet, men også ved foreldrene kunne påvirke behandlingsalliansen. Resultater fra moderator analyser kan bidra til at behandlingsprogrammene raffineres eller differensieres for bedre å kunne tilpasses variasjoner i målgruppens kjennetegn og problemer. Men det finnes også moderatoranalyser som viser at programmer kan være så robuste og fleksible at de virker i forhold til målgruppen varierer

betydelig med hensyn til sosial bakgrunn og individuelle kjennetegn (Borduin mfl., 1995; Ogden & Amlund-Hagen, 2009).

Oppsummert viser forskningen knyttet til evidensbasert behandling av alvorlige atferdsproblemer blant barn og unge at det finnes signifikante hovedeffekter av programmer når programbasert behandlingen har blitt sammenlignet med vanlige tiltak (eng. treatment as usual) i ordinær praksis. Selv om variasjoner i terapeut- og klientkjennetegn påvirker behandlingsutfall, så dokumenterer både primærstudier og meta-analyser klare programeffekter. Litteraturen som er referert i dette kapitlet viser at det er lite hensiktsmessig å sammenligne den programspesifikke tilnærmingen med fellesfaktor tilnærmingen uten å ta hensyn til hvilke problemer og målgrupper det dreier seg om. Mens det er dokumentert tydelige programeffekter i familie- og nærmiljøbasert behandling av alvorlige atferdsproblemer hos barn og unge, så er forskningsstatusen på andre problemområder og for andre målgrupper uavklart.

4. Randomiserte kontrollerte evalueringsstudier – det eksperimentelle designet.

Det eksperimentelle designet ligger til grunn for randomiserte kontrollerte evalueringsstudier (eng. Randomized Controlled Trials, RCT) og omtales ofte som 'gullstandarden' i intervensjonsforskningen (Cook & Campbell, 1979). Det skyldes at det bedre enn andre design egner seg for å teste om det er kausale sammenhenger mellom uavhengige og avhengige variabler, dvs mellom intervensjon og resultater. Designet innebærer sammenligning mellom ekvivalente grupper slik at en kan konstatere hvilket utbytte én gruppe har av en intervensjon sammenlignet med en tilsvarende gruppe som ikke får noe tilbud eller som får en annen intervensjon. Gjennom oppfølgingsstudier kan en også teste varigheten av klinisk endring.

I tillegg til å vise om intervensjoner fører til positive resultater kan en identifisere faktorer som moderer effektene av tiltaket, identifisere subgrupper som kan ha mer eller mindre utbytte enn gjennomsnittresultatene tyder på, samt identifisere og forklare mekanismene bak individuelle endringer. Det er med andre ord en naturlig progresjon mellom å undersøke om en intervensjon virker eller ikke virker, for deretter å fastslå for hvem og under hvilke betingelser den virker, samt hvordan den virker (Nock, Janis & Wedig, 2008).

I dette kapitlet presenteres og drøftes randomiserte kontrollerte design i forhold til evaluering av behandlingsprogrammer og andre intervensjoner. Planlegging av det idealtypiske eksperimentet må ta hensyn til forhold som (1) rekruttering og randomisering av deltakerne, (2) utvalgsstørrelse og statistisk styrke, (3) beskrivelse av kontrollbetingelsene, (4) måling av behandlingsintegritet, (5) valg av primære og sekundære resultatvariabler, (6) statistiske analysestrategier, (5) identifisering av potensielle moderatorer og mediatorer, (6) forholdet mellom statistisk og praktisk/klinisk signifikans og rapportering av effektstørrelse. Kapitlet avsluttes med en omtale av og kommentarer til the CONSORT statement (Moher mfl., 2001; 2010, Schulz, Altman & Moher, 2010).

Rekruttering og randomisering av deltakere

Rekruttering, inkludering og ekskludering av deltakere i randomiserte studier tar utgangspunkt i en klart definert målgruppe som behandlingen skal utformes og testes i forhold til og som er representative for populasjonen som en ønsker å generalisere resultatene til.

Inklusjons- og eksklusjonskriterier brukes for å forhåndsvurdere deltakerne slik at en får et definert utvalg som gjør det mulig å si noe om hvem behandlingen er nyttig for og hvem den ikke egner seg for. For å sikre et homogent utvalg kan en basere rekrutteringen på en bestemt diagnose som behandlingen retter seg mot og dermed redusere sannsynligheten for at tilleggsp problemer bidrar til å øke variasjonen i behandlingsutfall. En kan også identifisere målgruppen gjennom å benytte kriterieskårer i reliable og valide vurderingsskjemaer, eller intervjuer som fokuserer på målgruppens problemer. Individuer som skårer høyere en definert verdi kan være aktuelle deltakere i en klinisk behandlingsstudie.

Allerede i rekrutteringsfasen kan det skje en forhåndssiling av deltakere hvis de er selvrekrutterte (frivillige) heller enn ordinære klienter, og hvis deltakerne bare representerer de som aksepterer at det er en viss mulighet for at de kan havne i kontrollgruppen. Kunnskap om at det skal gjennomføres en randomisert evalueringsstudie kan også påvirke henvisningene til studien, hvis det er en mulighet for at klienten ikke får det behandlings-tilbudet som den henvisende instansen hadde som sitt førstevalg. Rekrutteringsmåten kan ha konsekvenser for resultatenes generaliserbarhet, og det anbefales derfor at forskerne i sine publikasjoner redegjør for inklusjons- og eksklusjonskriteriene (Schulz mfl., 2010).

Randomisering

Randomisering defineres av Moher mfl. (2010) på følgende måte: "With random allocation, each participant has a known probability of receiving each intervention before one is assigned, but the assigned intervention is determined by a chance process and cannot be predicted" (p.9). Tilfeldig fordeling eller randomisering av deltakere på ulike intervensjonsbetingelser eller behandlingsalternativer skal gi maksimal tillit til at det er intervensjonen heller enn noen annen påvirkning som er årsaken til resultatene (Rutter, Giller & Hagell, 1998). Randomiserte effektstudier har høy indre validitet, det vil si at en kan med høy grad av sannsynlighet kan fastslå kausalitet dvs. at det er en årsakssammenheng mellom intervensjon og resultater. Intervensjons- og kontrollgruppene skal være ekvivalente, dvs. at de er sammenlignbare med hensyn til alle kjennetegn som kan være relevant for resultatene, men med unntak av at den ene får en planlagt intervensjon som den andre gruppen ikke får (Rossi mfl., 2004).

I rapporteringen av randomiserte kontrollerte studier anbefales forskerne å beskrive hvordan deltakerne ble tilfeldig fordelt på intervensjons- og kontrollgruppen (Moher mfl., 2010) Eventuelle restriksjoner rapporteres, som for eksempel om utvalget var *stratifisert* dvs. at de før randomiseringen ble valgt ut etter visse kriterier (for eksempel et definert forholdstall

mellom gutter og jenter eller mellom familier med bestemte geografiske eller sosioøkonomiske kjennetegn). En bør også rapportere om randomiseringen ble foretatt av forskerne eller om den foregikk på hvert av utprøvningsstedene. Eventuelle stoppregler eller *blokking* bør også rapporteres, dvs. om randomiseringen opphørte når et bestemt antall deltakere var rekruttert (jf neste avsnitt om utvalgsstørrelse og statistisk styrke). Det bør også framgå hvem som registrerte deltakerne og hvem som laget den tilfeldige fordelingssekvensen. Randomiseringsprosedyren beskrives slik at det framgår om en benyttet nummererte containere (for eksempel i foreseglete konvolutter), myntkast eller brukte en datagenerert liste. Videre bør en rapportere eventuell *maskering*, eller skjult allokering, som innebærer at ingen vet hvilken gruppe deltakere trekkes til før inntakskartleggingen er gjennomført. Dette er ikke det samme som "blinding" som innebærer at man holder tilbake informasjon om hvilken behandlingsbetingelse deltakerne er fordelt på, overfor alle som kan bli påvirket av denne informasjonen. Mens maskering skal forebygge utvalgsskjevheter skal 'blinding' forebygge prestasjons- og vurderingsskjevheter. Og mens maskering kan implementeres i de fleste psykososiale behandlingsstudier, er 'blinding' vanskeligere å gjennomføre siden det er vanskelig å skjule hva slags behandling hver av gruppene får (Moher mfl., 2010; Shulz mfl., 2010).

Selv om randomisering i prinsippet skal føre til at en får ekvivalente grupper i hver av behandlingsbetingelsene, så kan det likevel oppstå utvalgsskjevheter, særlig hvis det dreier seg om små utvalg. Det er derfor vanlig å sammenligne gruppene på viktige bakgrunns- og resultatvariabler før intervensjonen iverksettes (eng. baseline comparisons) for å kontrollere at det ikke forekommer utilsiktede gruppeforskjeller som må korrigeres for i resultatanalysene. Gjennom en pretest kan en kontrollere om gruppene er sammenlignbare eller ekvivalente før en iverksetter tiltak, og dermed trekke riktige konklusjoner om hvor mye forandring som har forekommet i løpet av intervensjonen (Nock mfl., 2008). Moher mfl. (2010) anbefaler at en rapporterer demografiske og kliniske gruppeforskjeller ved inntak (basislinjen), men hvis en justerer for slike forskjeller i analysene, bør både justerte og ikke-justerte analyser rapporteres.

Prosedyrene for rekruttering og randomisering bør med andre ord være mest mulig åpne og transparente. Og når kravene til redegjørelse for rekruttering og randomisering av deltakere er såpass mange og detaljerte, er det for å unngå skjevrekuttering og for å sikre at gruppene er sammenlignbare. Men andre problemer kan oppstå, for eksempel hvis det er for få enheter å randomisere (som kan forekomme når analyseenheterne er skoler, kommuner eller land), eller

at en mangler enheter som kan være kontrollgruppe (for eksempel i prosjekter hvor alle deltakerne får intervensjonen). Randomisering kan også skape etiske eller juridiske problemer hvis noen mener at en holder tilbake behandling eller tjenester fra personer som har behov for dem (Asscher mfl. 2007; Lipsey & Cordray, 2000).

Statistisk styrke og effektstørrelse

Beregning av statistisk styrke inngår i arbeidet med å avgjøre hvor stort utvalg en trenger i en intervensjonsstudie for å kunne påvise signifikante behandlingseffekter. Statistisk styrke er sannsynligheten for at et estimat av programeffekten blir statistisk signifikant når den representerer en virkelig effekt av en viss størrelse. Den er en funksjon av effektstørrelsen som skal avdekkes, utvalgsstørrelsen, hva slags statistisk signifikanstest som anvendes og alpha nivået som settes for å kontrollere for type I feil (Cohen, 1992b; Rossi mfl. 2004). Når signifikansnivået vanligvis settes til 0.05, innebærer det at sjansene for at en pseudoeffekt (som kan skyldes 'statistisk støy') er like høy som den observerte programeffekten, er fem prosent eller mindre.

Forskeren må imidlertid selv estimere hvilken *effektstørrelse* som kan forventes, basert på tidligere studier eller på teoretiske forventninger. Beregninger av effektstørrelser bygger på statistikk som er utviklet i forbindelse med meta-analyser som "the standardized mean difference effect size", også omtalt som Cohens *d* (Cohen, 1988, 1992a, 1992b). Effekten uttrykkes da i standard avvik enheter og Cohens *d* beregnes som en brøk der telleren er differansen mellom middelerverdiene i de to gruppene og nevneren deres felles eller gjennomsnittlige standardavvik. Cohen (1992a) foreslo som konvensjon å skille mellom store effektstørrelser ($> .80$), medium effektstørrelse der forskjellen er stor nok til at den er synlig for en grundig observatør ($> .50$) og liten effektstørrelse som er mindre enn medium, men ikke så liten at den er triviell ($> .20$). Andre har foreslått at en effektstørrelse på $d = .20$ som standard for tilfredsstillende klinisk signifikans (Eyberg mfl., 2008) eller også så lavt som til $d = .10$ (Lipsey, 2000). Fordelen med effektstørrelser er at de ikke påvirkes av utvalgets størrelse, mens begrensningen er at de ikke gir informasjon om klinisk meningsfylte variasjoner i behandlingsrespons. En behandling som har en stor effekt på 20 prosent av klientene kan derfor framstå som mer effektiv enn en behandling som har mindre, men meningsfylt effekt på 90% av klientene (Westen mfl., 2004).

Metodelitteraturen anbefaler at statistisk styrke beregnes før undersøkelser gjennomføres, og påpeker at manglende statistisk signifikans kan skyldes at deltakergruppen er for liten til å avdekke klinisk meningsfylte forskjeller mellom behandlingsbetingelsene

(Moher, Dulberg & Wells, 1994; Rossi mfl. 2004). I intervensjonsforskning beregnes statistisk styrke primært for å estimere hvor mange deltakere en trenger i en studie for å avdekke statistisk signifikante og klinisk viktige forskjeller hvis slike finnes (Moher mfl., 2010).

Kontrollbetingelser

Det er den relative effektiviteten av planlagte intervensjoner som undersøkes i randomiserte studier; en måler hvor store endringene i målgruppen er sammenlignet med endringene i gruppen en sammenligner med. De vanligste kontrollbetingelsene i intervensjonsstudier består av ekvivalente grupper som ikke får noe tilbud, som får et minimalt tilbud eller som står på venteliste, eventuelt som får et annet aktivt behandlingstilbud. Gjennom valg av kontrollbetingelser avgjør forskerne hvor streng prøve de vil sette intervensjonen på, der sammenligninger med andre aktive intervensjoner utgjør det strengeste nivået (for eksempel når en sammenligner to empirisk validerte behandlingsprogrammer). Sammenligninger med ekvivalente grupper som ikke får noe tilbud, eller som står på venteliste, gir som regel bedre resultater enn sammenligninger med etablerte behandlingsalternativer. Nedenfor drøftes ulike kontrollbetingelser mer i detalj.

Sammenligningsgrupper som ikke får behandling. Den svakeste kontrollbetingelsen går ut på å sammenligne de som får behandling med noen som ikke får noen behandling. Kontrollgruppen får da kun den ekstra oppmerksomheten som følger med å bli kartlagt før og etter at intervensjonsgruppen har fått sin behandling. Men siden kontrollgruppen ikke blir lovet noen form for hjelp, er det alltid en mulighet for at de søker hjelp på annet hold. Det kan også reises etiske innvendinger mot ikke å gi noe tilbud til klienter som er hjelpsøkende for et potensielt alvorlig problem eller en lidelse. Et alternativ kan derfor være venteliste kontroll, der de som står på venteliste blir lovet behandling på et senere tidspunkt, som regel etter at forsøksgruppen har mottatt det planlagte tilbudet. Men venteliste kontroll har også sine begrensninger. Den eneste konklusjonen en kan trekke hvis en observerer signifikante forskjeller mellom eksperiment og venteliste kontroll gruppene er at å gjøre noe er bedre enn ikke å gjøre noe (Chambless & Hollon, 1998). En tester med andre ord ikke innholdet i behandlingen i forhold til alternativer. I venteliste design er det også vanskelig å fortolke effektstørrelser fordi validitetstrusler som forløp av tid, regresjon mot middelveiden, og spontan bedring i forhold til problemer med fluktuerende forløp, kan påvirke koeffisientenes størrelse (Westen mfl., 2004). Det er også en begrensning ved venteliste kontroll design at en

ikke kan studere langtidseffekter i forhold til en kontrollgruppe. Flere replikasjoner kan imidlertid styrke tilliten til resultatene. Hvis behandlingen virker, uansett av hvilken grunn, og hvis positive resultater replikeres i flere studier og med flere uavhengige grupper, så er behandlingen antagelig klinisk verdifull og nyttig.

Sammenligninger med andre former for behandling. Sammenligninger med andre aktive intervensjoner representerer den mest stringente testen på et behandlingsprogram. Da kontrollerer en for prosesser som er felles for alle former for behandling og som er uavhengig av behandlingen. En undersøker da om mekanismene bak en observert effekt handler om noe mer enn at klienten får oppmerksomhet fra en interessert person (terapeuten) eller at klienten får presentert forventninger om at han eller hun skal endre seg (Chambless & Hollon, 1998). Hvis kontrollgruppen får en behandling som er sammenlignbar med den som intervensjonsgruppen får, så blir det liten kontrast mellom gruppene (Kazdin, 1986). En slik situasjon kan oppstå hvis en sammenligner to virksomme programmer eller hvis ordinære tiltak holder en høy kvalitet (Löfholm mfl., 2009). Sammenligninger med ordinære tjenester kontrollerer imidlertid ikke for 'kompensatorisk rivalisering', noe som kan forekomme når terapeutene i sammenligningsgruppen anstrenger seg for å få bedre resultater enn de vanligvis gjør. De kan intensivere arbeidet med sine deltakere eller gi dem en positiv forskjellsbehandling i forhold til andre klienter de har (Cunningham, 2002). 'Programlekkasjer' kan også forekomme hvis forsøksintervensjonen ukontrollert smitter over på behandlingen som sammenligningsgruppen får. Kompensatorisk rivalisering og programlekkasje er eksempler på mekanismer som utilsiktet kan bidra til å redusere kontrasten mellom behandlingsbetingelsene. Lærdommen en kan trekke fra dette er at man må være mer oppmerksom på innholdet i behandling som sammenligningsgrupper får, enten deltakere mottar vanlige tjenester eller et annet aktivt tilbud. En bør derfor beskrive tilbudet til denne gruppen med hensyn til type, mengde og kvalitet like grundig som tilbudet til intervensjonsgruppen.

Ekvivalens. Av og til sammenligner en behandlinger med ukjent effekt med etablerte behandlinger for å demonstrere ekvivalens. Hvis en har grunn til å tro at behandlingstilbudet som allerede er etablert holder en tilfredsstillende kvalitet, kan en sammenligne dette med nye behandlingsformer for å teste om de er like effektive og holder samme kvalitet. Hvis en ny intervensjon ikke er signifikant dårligere enn den som er etablert og hvis den etablerte behandlingen heller ikke er bedre, så kan studien konkludere med at behandlingene er like effektive (Chambless & Hollon, 1998). Resultatene av slike sammenligninger kan imidlertid være vanskelige å fortolke. Blant annet må evaluatørene være sikre på at den etablerte behandlingen blir tilfredsstillende implementert og at det ikke dreier seg om en degradert

versjon av vanlig praksis (Klein, 1996). Det er derfor en fordel hvis noen på forhånd har evaluert hvor resultateffektive etablerte tilbud er før de sammenlignes med nye programmer eller tiltak.

For å oppsummere: Evaluering av behandlingsprogrammer forutsetter et sammenligningsgrunnlag og valg av kontrollgruppe påvirker generaliseringsmulighetene i en studie. Det er med andre ord viktig å gjøre rede for hva man sammenligner i randomiserte kontrollerte studier og være bevisst på hvor stringente prøver en setter intervensjoner på. For det første bør en kjenne til hvor effektiv gjeldende praksis er på et behandlingsområde, og for det andre bør en velge kontrollbetingelser ut fra hva som er praktisk gjennomførbart og etisk forsvarlig. I en kritisk kommentar hevder Westen mfl. (2005) at behandlingsforskere ofte foretrekker venteliste kontroll eller 'ingen behandling' fordi de gir større gjennomsnittlige behandlingseffekter enn sammenligninger med aktiv behandling. For å imøtegå slik kritikk bør en derfor sammenligne med et alternativt behandlingstilbud der hvor det er mulig. Et sammendrag av behandlingsprotokollen kan være gunstig samt en beskrivelse av opplæringen av terapeutene (Trudeau mfl., 2008).

Måling av behandlingsintegritet

Behandlingsintegritet handler om intervensjonens teorispesifikke metoder eller teknikker og ferdighetene som disse teknikkene og metodene blir implementert gjennom (Webb mfl. 2010). Måling av behandlingsintegritet er viktig for å kunne trekke valide konklusjoner om forholdet mellom behandling og resultater og handler om i hvilken grad intervensjonskomponentene blir iverksatt som angitt i behandlingsprotokollen. Behandlingsintegritet kan evalueres gjennom observasjoner, gjennom terapeutenes selvrapportering eller intervjuer med klienter (Perepletchikova mfl., 2009). Validiteten i slike evalueringer styrkes hvis kodere eller de som skal vurdere følger etablerte retningslinjer og en måler samsvaret eller reliabiliteten mellom koderne. En detaljert behandlingsprotokoll med beskrivelse av programmets kjernekomponenter kan gjøre det enklere å måle behandlingsintegritet og eventuelle avvik fra modellen. I tillegg kan integriteten styrkes gjennom en forhåndsbeskrevet opplæring av terapeutene og systematisk veiledning i tråd med intervensjonen. Samtidig kan det være vanskelig å vite hvordan man skal måle integritet på tvers av behandlingsformer; det som er viktige elementer i én behandling kan være mindre sentrale i andre behandlingstiltak.

Integritet er en viktig variasjonskilde eller moderator i forhold til hvilke resultater en oppnår i behandlingen. Selv om et behandlingsprogram er tilfredsstillende implementert kan

det være store forskjeller i behandlingsintegritet, dvs i hvor stor grad behandlingen som blir tilbudt den enkelte klient er i overensstemmelse med den opprinnelige modellen. Behandlingsintegriteten endres hvis terapeuten endrer dosering eller kvaliteten på behandlingen for eksempel på grunn av kjedsomhet eller økt erfaring og kompetanse (Nezu & Nezu, 2008). Forholdet mellom behandlingsintegritet og resultater uttrykkes ofte som en korrelasjon, men forskningen viser at høy grad av integritet (dvs. høy korrelasjon) ikke alltid gir de beste resultatene for klientene (Perepletchikova & Kazdin, 2005). En kan oppnå signifikante behandlingseffekter selv om behandlingsintegriteten er lav, ikke fordi behandlingen nødvendigvis er dårlig, men fordi den avviker fra behandlingsprotokollen. Klientene kan ha fått et bedre tilbud fordi terapeuten inkluderte flere virksomme teknikker, men som ikke hørte til metoden (Perepletchikova, Hilt, Chereji & Kazdin, 2009). Som nevnt i forrige kapittel er det snakk om en kompetent implementering av intervensjoner, noe som innebærer at en både kan legge til og trekke noe fra idealmodellen for å oppnå et bedre behandlingsresultat.

Perepletchikovas mfl. (2009) fant at få randomiserte studier rapporterte data om behandlingsintegritet og forklarte det med at arbeidet med å utvikle behandlingsprotokoller, gi grundig opplæring og veiledning av terapeuter, foreta video-opptak av behandlingssesjoner samt lære opp de som skal vurdere, kode og analysere data, krever en betydelig investering i form av penger, arbeid og tid. Samlet kan dette fungere som en barriere mot måling og rapportering av behandlingsintegritet. Manglende dokumentasjon gjør det vanskelig å estimere hvor viktig behandlingsintegritet er for resultatene i mange studier. Noen studier har gitt uventede resultater, noe en evaluering av et skolebasert rusforebyggende program for ungdom er et eksempel på (Sánchez mfl., 2007). Forskerne lyktes med å oppnå en høy implementeringsgrad, men denne var forbundet med økt heller enn redusert problematferd. Ungdommenes marijuana bruk økte i takt med undervisningsmengden og høyere frammøte til undervisningen predikerte økt tilknytningen til høyrisikovenner, alkoholbruk og sinne. Studien er en av flere som viser at gruppetiltak for høyrisiko ungdom øker sannsynligheten for 'gruppesmitte' og dermed høyrisiko atferd (Sánchez mfl., 2007; Dishion mfl., 1999; Dodge mfl. 2006). Men det finnes også forskningsstudier som viser at høy grad av behandlingsintegritet predikerer positive behandlingsresultater (Henggeler mfl., 1997; 1999; Mihalic & Irwin, 2003). Talsmenn for fellesfaktor tilnærmingen kritiserer studier av programeffekter for å behandle terapeuteffekter som feilkilder heller en varianskilder når de måler 'fidelity': "Fidelity measures attempt to control for the specific contributions of the therapist" (Blow, Sprenkle & Davis, 2007, s. 301).

Men selv om målinger av behandlingsintegritet er ressurskrevende og omdiskutert så har de en sentral plass i randomiserte evalueringsstudier for å dokumentere at det er en fullverdig versjon av en planlagt intervensjon man måler effekten av. Forskningen er imidlertid ikke entydig når det gjelder sammenhengen mellom høy behandlingsintegritet og positive behandlingsresultater; noen studier viser en positiv sammenheng, andre ikke (Mihalic mfl., 2004; Sánchez mfl., 2007; Sundell mfl., 2008; Waltz mfl., 1993; Webb mfl. 2010).

Resultatvariabler

Kontrollerte evalueringsstudier forutsetter at forskerne har hypoteser om hvilke resultatvariabler intervensjonen vil påvirke, hvilke endringer en kan forvente og hvordan en kan måle disse. Hvor aktuelle eller relevante ulike resultatvariabler er, avhenger av undersøkelsens målsetting og fokus. Resultatvariablene bør i noen grad være program-spesifikke og proksimale det vil si at de bør være tilpasset det enkelte program og fange opp de umiddelbare endringene i målgruppen etter avsluttet behandling, for eksempel reduksjon av utagerende atferd. Samtidig kan det være aktuelt å supplere med resultatmål som er generelle og langsiktige (distale), for eksempel forebygging av lovbrudd og fullføring av skolegang. Moher mfl. (2010) anbefaler at en i kontrollerte evalueringsstudier skiller mellom primære og sekundære resultatvariabler. De primære bør være de som behandlingen har hovedfokus på mens de sekundære kan omfatte både tilsiktede og utilsiktede resultater på andre relevante variabler. I behandlingen av alvorlige atferdsproblemer vil for eksempel de primære målene være knyttet til aggressiv og utagerende atferd mens sekundære mål kan være knyttet til internalisert problematferd eller skolefaglig fungering.

Ideelt sett bør en i intervensjonsstudier anvende multi-metode og multi-informant kartlegginger og det advares mot metode overlapping hvor en utelukkende baserer seg på informasjon fra en respondentgruppe (for eksempel selvrapportering) (Chambless & Hollon, 1998). *Multi-metode* kartlegging kan omfatte observasjon, atferdsvurderinger og intervjuer mens *multi-informant* kartlegging kan omfatte bidrag fra foreldre, lærere, barna selv og uavhengige observatører. Multi-informant og multi-metode kartlegging vil også som regel være *multi-setting*, dvs. bli gjennomført både hjemme og på skolen eller i barnehagen, eventuelt i framtiden. Gjennom å dekke flere arenaer kan en vurdere generalisering av atferdsendring fra hjem til skole eller omvendt. Noen mål eller variabler trenger man ikke reliabilitetsteste da de har høy 'face validity', så som antall arrestasjoner eller plasseringer utenfor hjemmet. Øvrige mål på avhengige og uavhengige variabler i randomiserte studier bør

ha vist seg å være reliable og valide i tidligere forskning (APA, 2001). I tillegg bør en vurdere instrumentenes *begrepsvaliditet*, dvs. om det måler begrepet på en dekkende måte og har tilfredsstillende psykometriske egenskaper.

Et viktig, men også problematisk tema er resultatvariablenes sensitivitet for forandring. Kartleggingsinstrumenter kan være sensitive mål på individuelle forskjeller på et gitt tidspunkt, men egner seg ikke nødvendigvis like godt til å fange opp endringer hos de samme individene over tid. Programspesifikke og endringssensitive mål vil bedre kunne registrere klinisk endring enn de som måler barn og unges generelle fungering og utvikling. Måleinstrumenter med høy sensitivitet for endring innebærer en stor mulighet for å finne signifikante behandlingseffekter på kort sikt, men det er usikkert hvor godt de fanger opp langsiktige resultater. Det har imidlertid vært lite fokus på og få analyser av hvordan en kan identifisere sensitive mål og hvordan sensitiviteten på kort og lang sikt kan økes (Lipsey & Cordray, 2000).

Oppsummert bør resultatmålinger omfatte flere metoder, flere informanter eller informasjonskilder og samles inn på flere arenaer som hjemme og på skolen. Gjennom å kombinere data fra flere informanter og flere metoder kan en motvirke effekten av rapporteringskjevheter (bias) og det forholdet at alle metoder har sine styrker og svakheter. Det er også en fordel å kombinere primære og proksimale resultatvariabler med mål som er sekundære og distale.

Statistiske analyser

Statistiske analysemetoder brukes for å sammenligne gruppene på primære og sekundære resultatvariabler men også for å gjennomføre tileggsanalyser av subgrupper og justerte analyser. Rapporteringen av slike analyser bør være så detaljerte at en kyndig leser med tilgang til originaldataene skal kunne replikere resultatberegningene (Moher mf., 2010). To viktige temaer i denne sammenhengen er spørsmålet om hvem som bør inkluderes i analysene og hvordan en skal håndtere innomgruppevariasjon.

Hvem bør inkluderes i analysene?

Et viktig tema i randomiserte studier er hvor mange som skal inkluderes i analysene, og hvordan en skal forholde seg til at noen avbryter behandlingen eller ikke møter opp i det hele tatt. Analysene kan enten omfatte alle som opprinnelig ble randomisert til intervensjons- og kontrollgrupper (intention-to-treat, ITT) eller bare de som har gjennomgått behandlingen.

Med uttrykket "intention-to-treat" (ITT) markerer en at analysene har blitt utført på grunnlag av hvilken behandlingsbetingelse deltakerne opprinnelig ble randomisert til, uavhengig av om de faktisk deltok i behandlingen eller ikke og uavhengig av om de fullførte hele eller bare deler av behandlingen. "Intention to treat" beholder med andre ord de som avslår eller som faller ut av studien i sine originale grupper. Dermed ivaretas randomiseringen og en opprettholder balansen mellom deltakernes kjente og ukjente kjennetegn på tvers av behandlingsbetingelsene. Analyser i henhold til "Treatment of the Treated" (TOT) inkluderer bare de som tilfredsstiller forskningsprotokollens krav til egnethet og til deltakelse i behandlingen og resultatevalueringen (omtales også som 'on treatment' eller 'per protocol'). Ulempen med bare å inkludere de som har fullført i analysene, er at frafallet i forsøks- og kontrollgruppen kan være selektivt, noe som igjen kan påvirke resultatene av analysene. Hvis for eksempel de mest problembelastede deltakerne har større tendens til å falle ut i én av gruppene, så vil resultatene kunne bli bedre for denne gruppen. Dermed kan resultatene bli en funksjon av at én av behandlingsbetingelsene bedre enn den andre lykkes i å holde på de vanskeligste. For å bøte på dette kan man enten analysere i henhold til ITT eller foreta en frafallsanalyse der en undersøker om frafallet er ikke-tilfeldig eller selektivt i noen av gruppene og om dette må tas hensyn til i vurderingen av resultatene.

Om en skal velge ITT eller TOT analyser beror på hensikten med studien. Effekstudier med lite frafall kan egne seg for TOT slik at man får testet hvordan behandlingen virker under optimale forhold. Effektivitetsstudier kan egne seg bedre for ITT analyser fordi de gir resultater som er typiske for situasjonen i ordinær praksis hvor det er vanligere at noen klienter ikke følger behandlingsopplegget eller trekker seg fra behandlingen (Wright & Sim, 2003). Både behandlingseffektivitet og behandlingseffekt er viktig, men kanskje for ulike formål: For klinikere kan det for eksempel være vanskelig å vurdere resultatene fra analyser som ikke tar hensyn til om deltakerne har fått den intenderte behandlingen eller om de har fått noen behandling i det hele tatt (Gibaldi & Sullivan, 1997; Wright & Sim, 2003). Andre kan være mer opptatt av de gjennomsnittlige resultatene en oppnår og der en tar hensyn til at ikke alle som får tilbudet gjennomfører behandlingen. I noen studier kan det være aktuelt å gjennomføre begge typer analyser, for eksempel i studier hvor det er et betydelig frafall i løpet av behandlingen, slik at det kan være store forskjeller mellom resultatene fra ITT og TOT analysene. I the CONSORT statement uttrykkes det eksplisitt at en gjennom å ekskludere deltakere fra analysene kan trekke feil konklusjoner. Så selv om ITT analyser kan føre til at en underestimerer potensialet i behandlingen så kan TOT ('per protocol') analyser gi misvisende resultater hvis frafallet ikke er tilfeldig (Moher mfl., 2010; Schulz mfl., 2010).

Konfidensintervall

I randomiserte studier uttrykkes hovedeffekter som kontrasten eller den gjennomsnittlige forskjellen mellom intervensjons- og kontrollgruppen. Et slikt enkelt punktestimat av behandlingseffekten kan imidlertid gi et lite dekkende bilde av variasjoner i individuelle endringer som reaksjon på en intervensjon. Noen endrer seg mye og andre endrer seg lite i løpet av behandlingen. Hvis en har tilstrekkelig statistisk styrke kan en få et godt statistisk estimat av den gjennomsnittlige program effekten, men estimatet gir lite informasjon om variasjonen rundt denne middelveien. For å bøte på svakhetene ved punktestimat kan en estimere analysens presisjonsnivå. En kan beregne gjennomsnittresultatene for hver av gruppene på hver av resultatvariablene samtidig som en estimerer konfidensintervallet som uttrykker usikkerhetsområdet for den 'sanne' behandlingseffekten. Et 95% konfidensintervall angir området der en forventer å finne den sanne verdien i 95 av 100 lignende studier (Moher mfl., 2010).

Komponent analyse

Eksperimentelle design behandler i utgangspunktet den uavhengige variabelen som en dikotom variabel; behandling versus kontroll, og intervensjonsprogrammet blir vurdert som en helhet (ofte omtalt som "black box"). Men innenfor programmet finner man flere variasjonskilder. Intervensjonen kan for eksempel være formidlet på varierende og ufullstendige måter og deltakerne kan reagere forskjellig på behandlingen eller på noen av dens enkeltkomponenter. Gjennom komponent analyser kan en sammenligne den relative effektiviteten av de ulike komponentene for å avgjøre hvilke komponenter som er nødvendige eller tilstrekkelige for optimal klinisk endring (Kaminski mfl., 2008; Nock mfl., 2008). Hvis en lykkes med å isolere de mest virksomme komponentene, kan en for eksempel forbedre behandlingsprogrammer gjennom å bruke mer tid og ha mer fokus på disse.

Moderatorer og analyser av subgrupper

Moderatorer er kjennetegn ved klient, terapeut, eller kontekst som påvirker forholdet mellom intervensjon og resultater, så som kjønn, etnisitet, sosio-økonomisk status eller tilleggs-problemer. Det dreier seg om variabler som vanligvis blir målt før intervensjonen, som er uavhengig av de eksperimentelle betingelsene, og som påvirker styrken eller retningen på relasjonen mellom intervensjonen og resultatet (Baron & Kenny, 1986). I behandling av

atferdsproblemer kan modererende kjennetegn ved barnet være debutalder for problematferd, problembelastning og komorbiditet, ved foreldrene kan det være sivilstatus, psykisk helse og konflikter, mens det i den sosiale konteksten kan dreie seg om kjennetegn ved skole, nabolag og kultur (Kazdin & Nock, 2003). En kan også inkludere teoretiske begrunnede moderatorer som forventninger og motivasjon i undersøkelsesdesignet (Nock mfl., 2008). Moderator analyser kan hindre at en trekker premature konklusjoner om programmers effektivitet basert på gjennomsnittlige hovedeffekter. Moderator analyser bør supplere heller enn erstatte analyser av hovedeffekter. De kan vise at et program med en positiv hovedeffekt er lite virksomt for noen undergrupper mens programmer som ikke har noen hovedeffekt kan likevel vise seg nyttig for visse undergrupper (Rossi mfl., 2004). Moderatoreffekter bør analyseres som interaksjonseffekter (for eksempel i ANCOVA og GLM analyser) og som analyser av "simple effects". Hvis resultatene varierer for to grupper, for eksempel gutter og jenter, eller barn over og under 8 år, kan både intervensjons- og sammenligningsgruppen deles inn etter kjønn eller alder og den gjennomsnittlige programeffekten for hver av kjønns- eller aldersgruppene kan sammenlignes.

En spesiell variant av moderator analyser tar for seg "*dose-response*" sammenhenger det vil si sammenhengen mellom intervensjons- eller behandlingsmengde og resultater. Her forutsetter en under ellers like forhold, at en større dose av behandlingen opp til et visst punkt skal føre til bedre resultater. Årsakshypotesen i randomiserte studier kan testes gjennom å undersøke om det er en dose-response sammenheng, dvs. om behandlingsutbyttet samvarierer systematisk med graden av intensitet, frekvens eller varighet av intervensjonen (Rutter mfl., 1998). I en systematisk kunnskapsoversikt over behandlingsstudier med barn og ungdom i perioden 1995-2004 sammenlignet Jensen, Weersing, Hoagwood og Goldman (2005) studier (n = 27) med ulike behandlingsnivåer, det vil si grupper som fikk den samme behandlingen, men med varierende dosering og intensitet. I 13 studier fant de en effekt av behandlingsnivåer eller dosering/intensitet mens i 14 studier gjorde de ikke det. Det kan imidlertid være et problem at behandlingsdosene ikke er tilfeldig fordelt. Hvis klientene som har de mest alvorlige og sammensatte vanskene får mest behandling, kan 'dose-response' analyser vise at de som får mest behandling får de dårligste resultatene. Dose-response sammenhenger bør derfor primært studeres i grupper med sammenlignbar problembelastning eller risikonivå (Rossi mfl. 2004).

Oppsummert kan moderatoranalyser bidra til å identifisere undergrupper som i ulik grad har utbytte av en intervensjon. Analyser av hovedeffekter kan kamuflere differensielle effekter for definerte subgrupper og kunnskap om forhold som moderer intervensjonseffekten

kan øke forståelsen av for hvem og under hvilke betingelser behandlingen er mest virksom. Men selv om interaksjonstester kan avdekke interessante behandlingseffekter i komplementære subgrupper, advarer the CONSORT statement mot å gjennomføre for mange subgruppe analyser fordi det øker muligheten for tilfeldige positive funn (Moher mf., 2010).

Endringsmekanismer – mediering

Randomiserte behandlingseffektstudier har blitt kritisert for ikke å kunne gi informasjon om *hvordan* behandlingen virker dvs. i liten grad kunne belyse hva som bidrar til endring (Kazdin & Nock, 2003). Når en rapporterer hvor mye behandlingsbetingelsene forklarer av variansen i resultatvariablene, så sier ikke det noe om hvilke endringsmekanismer som er virksomme. Med endringsmekanismer menes prosesser eller hendelser som leder til og forårsaker endring og som forklarer hvorfor og hvordan en behandling virker. Dette omtales også som mediering og påvises gjennom fire logiske relasjoner mellom behandling, mediator og resultater (Baron & Kenny, 1986): (1) behandlingen gir forventede resultater, (2) de terapeutiske intervensjonene påvirker bestemte mekanismer, (3) de terapeutiske mekanismene påvirker graden av psykopatologi og (4) de terapeutiske effektene forklares gjennom denne årsaksstien.

I en gjennomgang av 67 kliniske behandlingseffektstudier fant Weersing og Weisz (2002) at bare 6 av dem analyserte mediering. De omtalte dette som terapeutiske mekanismer og fant at de oftest ble studert i foreldre- og systemrettede intervensjoner som har en sentral plass i behandlingen av barn og unge med alvorlige atferdsproblemer. Forfatterne lokaliserte 22 behandlingseffektstudier av PMT (Parent Management Training) og MST (Multisystemic Therapy). Studiene av PMT tydet på at endringer i foreldrenes oppdragslesferdigheter var en virksom terapeutisk mekanisme. I MST fant de indikasjoner på at behandlingseffekt ble mediert av økt foreldretilsyn med ungdommen, forbedringer av familierelasjoner (familiefungering, samhold og varme) og redusert kontakt med avvikende venner (Huey mfl. 2000). Weersing og Weisz (2002) påpeker imidlertid at det var vanskelig å bekrefte at de teoretisk funderte endringsmekanismene var virksomme fordi de fleste studiene målte mediatoren ved slutten av behandlingen. Hvis resultatvariablene og medierende variabler måles samtidig kan årsaksretningen gå begge veier: det er for eksempel lettere for foreldre å ha tilsyn med ungdom hvis disse slutter å være sammen med avvikende venner, men økt foreldretilsyn kan også føre til redusert kontakt med avvikende venner. Hvis en skal fastslå at endringer i mediatoren fører til endringer i atferd, er det imidlertid nødvendig å måle endringer i mediatoren på et tidligere tidspunkt enn endringene i resultatvariabelen. En annen

svakhhet ved medieringsanalyser kan være metodisk overlapping som forekommer når den samme informanten rapporterer både medierings- og resultatvariabelen.

I den tidligere omtalte kunnskapsoversikt undersøkte Jensen, Weersing, Hoagwood og Goldman (2005) i hvilken grad behandlingsstudier med barn og ungdom (i perioden 1995-2004) inneholdt analyser av underliggende endringsmekanismer. Oversikten omfattet 52 RCT studier med to eller flere behandlingsbetingelser og det var få av dem som rettet seg mot eksternalisert problematferd. Blant disse var det 29 studier som sammenlignet en antatt virksom behandling med én eller flere tilsvarende intensive behandlinger med andre terapeutiske ingredienser. I 14 studier fant de gruppeforskjeller mens i 15 fant de ingen forskjeller. Det vil si at bare halvparten av eksperimentgruppene hadde bedre resultater enn kontrollgruppene. I begge gruppene kunne få av studiene fastslå om det var de antatt aktive terapeutiske ingrediensene som forklarte graden av endring. Studiene rapporterte heller ikke om man hadde testet alternative forklaringer, som positive forventninger, terapeutisk allianse eller oppmerksomhet. Oppsummert tydet funnene på at mange studier av psykoterapi med barn og unge ikke kontrollerte for terapeutiske fellesfaktorer og inneholdt heller ikke analyser av spesifikke endringsmekanismer (mediatorer). Bare 5 av 52 studier inneholdt medieringsanalyser. Blant disse var Eddy og Chamberlains (2000) studie av Multidimensional Treatment Foster Care (MTFC) der analysene viste at effekten av behandlingen ble mediert gjennom økt foreldretilsyn, positive relasjoner mellom voksne og ungdom, og redusert kontakt med negative venner. Kombinasjonen av 'family management skills' og kontakter med avvikende venner forklarte 32% av variansen i senere antisosial atferd.

Jensen mfl. (2005) konkluderte med at det er en stor mangel på kunnskap om "aktive ingredienser" i form av spesifikke og uspesifikke terapeutiske faktorer som underbygger endringer i barn og unges atferd. Forfatterne mente at en i for liten grad har studert psykologiske endringsmekanismer i undersøkelser om atferdsendring (som for eksempel sosial kognisjon, motivasjon, self-efficacy og forventninger). De anbefalte derfor å legge mer vekt på analyser av terapeutiske fellesfaktorer som oppmerksomhet, positiv anerkjennelse og terapeutisk allianse eventuelt på effekten av behandlingsintensitet og dosering eller andre prosesser som kan mediere terapeutisk forandring (Jensen mfl., 2005).

På samme måte som korrelasjoner ikke påviser årsaksforhold, så beviser heller ikke mediering at det er snakk om en endringsmekanisme. Stringente kriterier for å påvise en endringsmekanisme i psykososiale intervensjoner er ifølge Kazdin & Nock (2003): (1) en sterk forbindelse i form av en signifikant korrelasjon mellom behandling (T), en antatt mekanisme (M) og et resultat (O), (2) spesifisitet der en viser at T er unikt relatert til endring i

M, og at endring i M er unikt relatert til endring i O, (3) en gradient som viser at det er en dose-respons sammenheng mellom T, M og O, (4) en temporær sammenheng som viser at endringene i M forekommer tidsmessig før endringene i O, (5) konsistens som viser replikasjon av observerte resultater på tvers av utvalg, forskere og settinger, (6) eksperimentelt design som RCT, (7) plausibilitet som innebærer at funnene må være troverdige og rimelige og (8) logisk sammenheng mellom funnene og den vitenskaplige kunnskapen en ellers har på området. Selv om det neppe finnes studier som tilfredsstiller alle kriteriene, så finnes det eksempler som kommer tett opp til dette. Kazdin og Nock (2003) viser også (som Weersing og Weisz, 2002) til Huey mfl., (2000) som fant at terapeutenes adherence til MST protokollen var forbundet med bedre familierelasjoner og redusert kontakt med avvikende venner, og dette var igjen forbundet med reduksjoner i normbrytende atferd.

Det finnes et stort antall psykoterapier (Kazdin, 1997), og det er sannsynlig at mange av dem har felles endringsmekanismer. Gjennom å identifisere endringsmekanismer kan en unngå overlapping mellom behandlingsmodeller og en kan øke behandlingsutbyttet for klienter, fordi en får kunnskap om nødvendige, tilstrekkelige og fremmede ingredienser og innenfor hvilke behandlingsmengder (doser) disse er virksomme (Kazdin & Nock, 2003). Hvis en kan identifisere mekanismene som føre til at individer endrer seg som en følge av psykososiale intervensjoner, så vil en også kunne belyse prosesser som mer allment fører til atferdsendring (Nock mfl., 2008).

Statistisk og praktisk-klinisk signifikans

Det første trinnet i analyser av behandlingseffekt består som regel av statistisk signifikantesting, for eksempel med ANCOVA analyser. I noen tilfeller kan det være en statistisk signifikant forskjell mellom behandlingsbetingelsene selv om det faktisk ikke er noen program effekt (type I feil), men det motsatte kan også forekomme; at en ikke finner noen statistisk signifikant forskjell selv om programmet faktisk er virksomt (type II feil). Type I feil, dvs. å påvise statistisk signifikans når det ikke er noen program effekt, kan kontrolleres gjennom å høyne signifikansnivået. Å kontrollere for type II feil er imidlertid vanskeligere fordi det forutsetter at forskningsprosjektet har tilstrekkelig statistisk styrke (jf tidligere avsnitt).

Statistisk signifikans har blitt kritisert for å være en dårlig eller misvisende standard fordi den rapporterer lite annet enn at forskjellene mellom behandlingsbetingelsene antagelig ikke skyldes tilfeldigheter (Cohen, 1994; Lipsey, 1999). Som tidligere omtalt kan en imidlertid med små justeringer i et eksperimentelt design teste moderatorer og mekanismer

for klinisk endring (Cook, 2003). Men i tillegg til å være statistisk signifikant bør behandlingseffekter også være klinisk eller praktisk meningsfylte. Kazdin (2008) minner for eksempel om at statistisk signifikans påvirkes av utvalgets størrelse, og hvor mye variasjon det er mellom individene. Manglende signifikans kan blant annet skyldes at utvalget er for lite og at man får for lite statistisk styrke til å påvise reelle forskjeller mellom behandlingsalternativer. Motsatt kan en også gjennom å benytte store utvalg få relativt trivielle forskjeller i behandlingsutfall til å framstå som signifikante. Så selv om en liten effekt teoretisk sett kan være så robust at den ikke kan forklares av tilfeldigheter, så er det ikke sikkert at den er klinisk interessant eller verdifull (Chambless & Hollon, 1998). De tre vanligste indikasjoner på klinisk signifikans er i følge Kazdin (2008): (1) at høye symptomskårer ved starten av behandlingen faller innenfor det normative området ved slutten av behandlingen, (2) at store endringer forekommer hos individer i løpet av behandlingen (for eksempel en problemreduksjon på 30 prosent eller et standardavvik), og (3) at individet ikke lenger innfrir kriteriene for en psykiatrisk diagnose. Noen anbefaler praktisk-kliniske indikatorer som prosentvis hvor mange av deltakerne som forbedret seg, hvor mange som ble dårligere og hvor mange som ikke viste tegn på forandring etter behandlingen (Lipsey & Cordray, 2000; Speer & Greenbaum, 1995). Andre kriterier for praktisk-klinisk signifikans kan være at individene oppnår personlige eller samfunnsmessig viktige mål som å kunne bo i et fellesskap med minimum av tilsyn eller kan holde på en jobb. Kazdin (2008) minner også om at mål som statistisk signifikans, klinisk signifikans eller effektstørrelser sier lite om i hvilken grad klienten og familien har fått hjelp i sitt daglige liv. I tillegg er den kliniske nyttedimensjonen et spørsmål om behandlingstilbudet er lett tilgjengelig og kostnadmessig konkurransedyktig i forhold til andre tiltak (Glasgow mfl., 2003).

Oppfølgingsstudier

Spørsmålet om når man skal måle resultater er like viktig som hvordan. Den umiddelbare responsen på behandlingen er noe annet enn vedlikeholdseffekter og oppfølgingsstudier er derfor viktige for å undersøke om endringene i atferd og reaksjoner varer over tid. Resultater fra oppfølgingsstudier kan være vanskelig å fortolke hvis en ikke har kontroll på det som skjer med deltakerne i oppfølgingsperioden og resultatene påvirkes av selektivt frafall i undersøkelsesgruppene. Behandling kan ha forsinket effekt på sosial fungering og symptomer, og kan vise seg først lenge etter at behandlingen er avsluttet. Det står klientene fritt å søke ytterligere behandling hvis de ønsker det; de kan søke behandling fordi symptomer eller problemer kommer tilbake, men også hvis nye problemer oppstår. Et annet problem med

oppfølgingsstudier er at de på tilsvarende måte som pre-post studier er sårbare for resultatskjevheter som skyldes *differensielt frafall*. Som tidligere nevnt kan et behandlingsprogram framstå som mer effektivt enn kontrollbetingelsene fordi det lykkes dårligere med å holde på høyrisiko klienter (Klein, 1996). Det er imidlertid ikke klart hvordan en skal skille mellom klienter som avbryter behandlingen og klienter som ikke møter opp til behandlingen i det hele tatt. Begge grupper representerer uønskete frafall, men manglende oppmøte og frafall kan ha forskjellige årsaker og konsekvenser. Det finnes ingen enkle løsninger på dette problemet, men Chambless & Hollon (1998) anbefaler at en samler inn dataene for alle (dvs. ITT analyser), men at en tolker dem med forsiktighet.

Det er heller ikke klart hvor lenge en bør gjøre oppfølgingsstudier. Flay mfl. (2005) anbefaler at det bør foreligge minst én signifikant langtidsoppfølging av forebyggende programmer som er 'efficacious', og antyder en oppfølgingsperiode på seks måneder mens Blueprint serien anbefaler oppfølging etter et år (Elliott, 1998). Hvor lenge en skal fortsette oppfølgingen bør avhenge av det naturlige forløpet av det aktuelle problemet, samt styrken og stabiliteten i behandlingseffekten (Chambless & Hollon, 1998). Siden mange av de familie- og foreldretreningsprogrammer rettet mot alvorlig atferdsproblemer hos barn og unge har som målsetting å forebygge alvorlig antisosial atferd i ungdoms- og voksen alder, er det viktig å kunne dokumentere slike langtidseffekter. Slik dokumentasjon foreligger imidlertid i liten grad og må sies å være en svakhet ved mange av studiene av tidlig intervensjon og forebygging av alvorlige atferdsproblemer. Ett av unntakene her er signifikante oppfølgingsresultater for Multisystemisk Terapi i USA etter fire (Borduin mfl., 1995) og 13,7 år (Schaeffer & Borduin, 2005). Det knytter seg både økonomiske, praktiske og etiske problemer til slike langtidsoppfølginger, da det forutsetter langsiktig finansiering, en stabil forskergruppe og at deltakerne samtykker i å bli kontaktet til dels mange år etter avsluttet behandling. Det mest realistiske er derfor antagelig å studere langtidseffekter gjennom registerdata.

Rapportering av randomiserte kontrollerte studier

Hvis lesere av publikasjoner om randomiserte kontrollerte behandlingseffektstudier skal kunne forstå hvorfor og hvordan studiene ble gjennomført og data ble analysert, er det viktig at rapporteringene er åpne, transparente og til dels standardiserte. En autoritativ veiledning i hvordan en bør rapportere slike studier ble utarbeidet av en gruppe kliniske forskere, statistikere, epidemiologer og redaktører og omtales som the CONSORT statement

(Consolidated Standards for Reporting Trials, Moher mfl., 2001, 2010; Schulz mfl., 2010). I en oppdatering av CONSORT statement (Moher mfl., 2010; Schulz mfl., 2010) ble hovedinnholdet fra 2001 beholdt, men det ble foretatt visse forenklinger og noen klargjøringer av kriterienes ordlyd. Videre ble noen kriterier flyttet og noen ble delt opp i underpunkter. Hensikten var fremdeles å sikre presisjon i analysene, blant annet at analysene i tester av hovedeffekter må være på samme nivå som randomiseringen og omfatte alle deltakerne (m.a.o. ITT). Den siste versjonen av the CONSORT statement inneholder 25 punkter og underpunkter med overskriftene: tittel og sammendrag, introduksjon, metode, resultater, diskusjon og annen informasjon (om registrering, protokoll og finansiering) (Schulz mfl., 2010). Mangelfull rapportering i forhold til disse kriteriene kan i følge forfatterne føre til skjeve eller feilaktige (biased) estimater av behandlingseffekt og informasjonen er avgjørende for å kunne bedømme forskningsfunnenes reliabilitet og relevans. Kriteriene i the CONSORT statement har hatt stor innflytelse på forskningstidsskrifters publiseringskrav til randomiserte kontrollerte studier, og har utvilsomt også bidratt til å heve studienes rapporteringskvalitet. Flere av punktene i oversikten nedenfor er omtalt tidligere i dette kapitlet.

I tillegg til rapporteringskriteriene inneholder the CONSORT statement utkast til et flytdiagram som viser deltakerflyt gjennom alle faser av et forskningsprosjektet, dvs. hvor mange som ble rekruttert til studien og hvor mange som fullførte i hver av behandlings-betingelsene. For hver av gruppene (intervensjons- og sammenligningsgruppe) rapporteres det hvor mange deltakere som ble tilfeldig fordelt, hvor mange som fikk den planlagte behandlingen, hvor mange som fullførte behandlingen og som ble analysert i forhold til de primære resultatvariablene (Schulz mfl., 2010).

The CONSORT statement har primært sitt utspring i klinisk medisinsk forskning og har også hentet de fleste av sine eksempler derfra (Moher mfl., 2010). Det kan hevdes at kriteriene for rapportering er bedre tilpasset testing av medikamenter enn testing av omfattende behandlingsprogrammer, og en svakhet kan sies å være at de fleste kriteriene har til hensikt å styrke undersøkelsenes indre validiteten. Det betyr at det kan være behov for å supplere med kriterier som bidrar til å belyse resultatenes generaliseringsverdi og dermed den ytre validiteten. Glasgow og kolleger (1999, 2003) hevdet at sammenlignet med medisinske intervensjoner er atferdsintervensjoner ofte vanskeligere å definere og standardisere. Slike intervensjoner bør blant annet appellere til en bredt definert populasjon eller målgruppe, inkludert de som ikke er selvrekrutterte. Både i effekt og effektivitetsstudier er det viktig å beskrive hvilke intervensjoner som kan: nå en bredt definert målgruppe, bli tatt i bruk av flere organisasjoner/ etater som skoler og helsetjenester, bli implementert av et bredt sammensatt

personale med varierende grad av opplæring og ekspertise, og føre til replikerbare og langvarige positive resultater (og minimalt negative resultater) til en rimelig kostnad. Dette stiller bestemte krav til statistiske tilnærminger som kan håndtere individuelle så vel som kontekstuelle faktorer (dvs. flernivå analyser – multi-level). Glasgow og kolleger (1999; 2003) oppsummerer sin modell i akronymet RE-AIM som står for reach (R), efficacy/effectiveness (E), adoption (A), implementation (I), og Maintenance (M). Med denne modellen vil de understreke at randomiserte studier ikke bare skal formidle relevante endringer på individnivå, men også på intervensjons- eller miljønivå gjennom å formidle hvor stor andel av målgruppen de når fram til og hvor generaliserbart intervensjonen er utover forskningsperioden.

Siden nesten ingen av kriteriene i the CONSORT statement (Moher mfl., 2001) dreide seg om generalisering og ytre validitet foreslo Glasgow mfl. (2003) følgende syv tilleggskriterier: (1) angi hvilken målgruppe som studien skal generaliseres til, (2) rapportér andelen som blir ekskludert, deltakelsesfrekvensen blant dem som kvalifiserer for intervensjonen og deres representativitet, (3) rapportér hvilke metoder som ble brukt for å rekruttere utprøvningssteder, inkludert hvor mange som ble ekskludert, deltakerraten blant de som ble forespurt, og hvor representativitet utprøvningsstedene var, (4) angi populasjonen av tjenesteytere (eng. intervention agents) som man ønsker skal implementere programmet og hvor representative de som deltar i undersøkelsen er for de som senere skal levere tjenestene, (5) rapportér i hvilken grad ulike intervensjonskomponenter blir levert (av ulike intervensjons-agenter) som angitt i protokollen, (6) rapportér hvor mye tid og hvilke kostnader som kreves for å tilby intervensjonen, og (7) rapportér kontinuitet eller diskontinuitet av intervensjonen på organisasjonsnivå og hvordan den eventuelt tilpasses eller modifiseres når undersøkelsen er avsluttet. I tillegg bør en gjennom oppfølgingsstudier undersøke hvor godt resultatene vedlikeholdes på individnivå. Glasgow mfl. (2003) fokuserte med dette på hvordan virksomme intervensjoner adopteres og fungerer i organisasjoner og blant praktikere som skal anvende dem i ordinær praksis. De var opptatt av hvor godt intervensjonen nådde ut til målgruppene, men også i hvilken grad de blir videreført i praksismiljøer. Vedlikehold av programmet er både et spørsmål om vedlikehold av atferdsendringer i målgruppen og vedlikehold av programtilbudet i aktuelle praksismiljøer.

Oppsummerende diskusjon

Det finnes ingen gode alternativer til det eksperimentelle designet hvis en vil besvare spørsmålet om programmer faktisk virker. Et sentralt tema i slike studier er avveiningen mellom hensynet til indre og ytre validitet. Hvis en ønsker å maksimere den indre validiteten er det viktig å ha homogene deltakergrupper og standardiserte intervensjoner som i effektstudier. Slike deltakere og behandlingsbetingelser er imidlertid lite representative for ordinær praksis, og generaliseringsmulighetene øker dersom en har få eksklusjonskriterier og dersom deltakere har vanlig forekommende tilleggsproblemer, som i effektivitetsstudier. Randomiserte studier har ifølge Westen mfl. (2005) to funksjoner. For det første bidrar de til en klinisk og vitenskaplig identifisering av virksomme intervensjoner og for det andre brukes de for å validere manualiserte behandlingspakker som kan formidles til klinisk praksis. De stiller seg mer positive til den første funksjonen der hensikten er å lære om intervensjoner, strategier, teoretiske tilnærminger, prosessvariabler, eller klient- og terapeut kjennetegn som kan være nyttige eller virksomme i psykologisk behandling. Mer kritiske er de til en pragmatisk og politisk bruk av randomiserte studier for å vise at korttids psykologisk behandling er like effektiv som medisinsk behandling. Fordi RCT ideelt sett forutsetter homogene utvalg av deltakere mener de at det er uklart hva man skal gjøre med klienter som i ordinær praksis viser symptomer som ikke passer inn i kriteriene. Innvendingen fra Westen mfl. (2005) er ikke uforenlig med utvikling og formidling av evidensbaserte behandlingsprogrammer og de senere års forskning har vist at randomiserte kontrollerte studier av evidensbaserte intervensjoner kan vise til positive hovedeffekter og ta høyde for de vanligste tilleggsproblemer som klientene har.

Weisz, Weersing og Henggeler (2005) imøtegår også denne kritikken ved først å si seg enige i at randomiserte design har begrensninger og at det finnes alternative måter søke kunnskap på. De vedgår også at rapportering av resultater i slike design bør være mer utfyllende og mer sammenfallende på tvers av studier. Deretter hevder de at selv om denne metodologien ikke er perfekt, så betyr det ikke at den bør avskrives til fordel for andre tilnærminger. Psykoterapiforskningen i betydningen rigorøs behandlingsforskning i vanlig praksis er etter deres oppfatning bare i startfasen. I denne fasen vil de fleste behandlingsformer og deres forskningsgrunnlag ha begrensninger som gjør dem lette å kritisere. De etterlyser derfor nyttig kritikk der en på grunnlag av en korrekt beskrivelse, identifiserer årsakene til problemene og foreslå forbedringer. Utviklingen av 'best practices' forutsetter

etter deres mening at forskning og praksis samarbeider for å utvikle intervensjoner som både innfrir kravene til evidens, praktisk nytte (relevans) og effektivitet (Weisz mfl. 2005).

Rapporteringskravene i den idealtypiske RCT er omfattende, noe som framgår av the CONSORT statement og forslag til tilleggskriterier. Randomiserte studier i intervensjonsforskning har tradisjonelt hatt fokus på den relative effekten av planlagte intervensjoner i forhold til kontrollbetingelser som kan være ingen behandling, venteliste kontroll eller sammenligninger med andre aktive behandlingsformer. Men designet har også blitt modifisert og utvidet slik at det i økende grad retter fokus mot forhold som modererer behandlingseffekten som for eksempel alder og kjønn og forhold eller prosesser som medierer endring, så som familiesamhold og foreldreferdigheter.

5. Parent Management Training, Oregon og Multisystemisk Terapi – programkjennetegn og forskning

Flere strukturerte program for behandling av alvorlige atferdsproblemer blant barn og unge er utviklet de siste årene, særlig i USA, og flere av dem er implementert i Norge, blant annet ”De utrolige årene” (Larsson mfl., 2009; Webster-Stratton & Herbert, 1994), Parent Management Training (Patterson & Forgatch, 1995) og Multisystemisk terapi (Henggeler m.fl., 1998, 2009). Som nevnt er de to sistnevnte programmene tema for denne avhandlingen og vil bli nærmere beskrevet i dette kapitlet. Parent Management Training, Oregon modellen (PMTO) retter seg mot foreldrene til barn med alvorlige atferdsproblemer i alderen 12 år eller yngre mens MST har en mer markert systemisk eller økologisk profil og retter behandlingsinnsatsen mot aldersgruppen 13-17 år, deres familie og sosiale nettverk.

Parent Management Training – Oregonmodellen (PMTO)

Parent Management Training inngår i gruppen ”foreldretreningsprogram” (parent training programs) som bygger på antakelsen om at manglende eller uegnede foreldreferdigheter bidrar til at barns atferdsproblemer oppstår, utvikler seg og opprettholdes (Lundahl, Risser & Lovejoy, 2006). Foreldreprogrammer finnes i mange varianter med forskjellig behandlingsomfang og med ulike krav til terapeutenes kvalifikasjoner. De varierer også med hensyn til formidlingsform (individuell, gruppe eller selvinstruksjon), terapeutiske komponenter og målgruppe (foreldre hver for seg eller sammen). Atferdsorienterte foreldreprogrammer har som mål å endre barns atferd gjennom å motivere og påvirke foreldre til å endre sin forståelse, oppfatninger, kommunikasjon og interaksjon med barnet. De har vist seg å kunne skape endringer i de sosiale betingelsene i familien slik at barns prososiale atferd blir bekreftet av foreldrene mens deres negative atferd blir etterfulgt av negative konsekvenser (straffet) eller ignorert (Serketich & Dumas, 1996). Et av de eldste og best etablerte programmene er i følge Brestan og Eyberg (1998) Oregon modellen av Parent Management Training (PMTO, Patterson, Reid, Jones & Conger, 1975; Patterson, 1982, 1986, 2005).

PMTO: teori, behandlingsprogram og forskning

Gerald Pattersons teoretiske modell for sosial læring og sosial interaksjon forklarer hvordan barn lærer antisosial atferd gjennom tidlige erfaringer i familien, og hvordan slik atferd kan videreutvikles i kontakt med andre personer i barnets omgivelser som venner, medelever og lærere (Patterson, 2002; 1982). Antisosial atferd utvikles i følge modellen gjennom en gjensidig læringsprosess mellom foreldre og barn, og mangelfulle eller uegnede foreldreferdigheter er den viktigste drivkraften bak barns avvikende atferd (Patterson mfl., 1992). Siden antisosial atferd er lært, kan den også avlæres eller forebygges gjennom læringsbasert behandling av foreldre og barn, dvs. gjennom 'parent management training'.

Patterson om utvikling av antisosial atferd

Pattersons (1982) teoretiske modell for sosial interaksjon og sosial læring presenteres som en handlingsteori og en målingsteori. Den legger vekt på å forklare hvorfor barn *praktiserer* antisosial atferd like mye som hvordan de *lærer* den og stiller krav om at alle elementene i modellen må kunne måles på en reliabel måte (Patterson, 1986). Innen de er fem år gamle lærer de fleste barn seg mange former for aggressiv atferd, men etter hvert forstår de at de oppnår mer gjennom prososial atferd, og den aggressive atferden avtar (Patterson, 1982). En forutsetning for at de skal utvikle prososial atferd er at de opplever foreldrenes reaksjoner som forutsigbare eller betingede slik at positiv atferd bekreftes og belønnes mens negativ atferd møtes med negative reaksjoner. De fleste barn lærer på denne måten at det er mer lønnsomt å være prososial enn antisosial. Men noen barn fortsetter å utagere for å nå sine mål, og for disse ser det ut til at negativ eller aversiv atferd gir flere fordeler enn prososial atferd. Disse barna kan ha bestemte nevrologiske og fysiologiske kjennetegn som sammen med et vanskelig temperament gjør dem vanskelige å påvirke eller kontrollere. De er ofte mer aggressive enn det som er vanlig for alderstrinnet og tidlig kan de vise tegn på motorisk uro og impulsivitet slik at de har vanskelig for tilpasse seg regler og akseptere voksenautoritet. Men drivkraften i utviklingen av antisosial atferd er i følge Patterson (1982, 2005) tvingende samspill i familien (eng. the coercive family process).

"Coercion theory" - tvingende samspill i familien. Tvingende samspill (eng. coercion) kan beskrives som en fastlåst og tvangspreget familieprosess med hyppige og opptroppende foreldre-barn konflikter (Patterson, 2002; Solholm mfl., 2005). «Coercion»-begrepet beskriver hvordan barn og foreldre gjensidig former hverandres atferd og kontrollerer hverandre gjennom aggressive og aversive handlinger. Foreldrene opplever at barnet ikke

samarbeider eller at det er ulydig mens barnet på sin side lærer at foreldrene ofte gir seg når de utsettes for motangrep i form av høylydte protester eller raserianfall. På denne måten lærer noen barn hvordan de kan unngå ulystbetonte oppgaver eller plikter, eller hvordan de kan mase eller true til seg noe de vil ha. Konfronterende samhandlingssekvenser gjentar seg regelmessig og øker sannsynligheten for gjensidig opptrapping av aggresjonen mellom familiemedlemmene, og for at konfliktene forlenges og utvides. Noen ganger lykkes foreldrene med å få barnet til å gjøre som de ber om, men ofte er det barnet som får viljen sin, noe som over tid kan føre til at foreldrene resignerer og gir etter (Patterson og Forgatch, 2000). Samhandling preget av 'tvingende samspill' kan føre til at foreldrene unngår konflikter gjennom umiddelbart å gi seg, eller at de svarer med motaggresjon i håp om å gjenvinne kontroll over barnet. Resultatet kan også være at foreldre og barn forsøker å unngå hverandre slik at de gjør få ting sammen. De spiser for eksempel til forskjellige tider og oppholder seg sjelden i samme rom. Hjemme svekkes foreldrenes tilsyn med barnet hvis de ikke følger opp daglige rutiner eller følger med på hva barnet foretar seg. Fra cirka 10-årsalderen oppholder også barn seg stadig mer utenfor hjemmet, og påvirkningen fra jevnaldrende og skolemiljø blir mer merkbare. Dette kan føre til ytterligere svikt i foreldrenes tilsyn slik at de sjelden vet hvor barnet er, hvem det er sammen med, hva det holder på med eller når det kommer hjem (Patterson, 1982; 1986). Dette øker mulighetene for at barnet har kontakt med venner og miljøer der de har muligheter for å observere og delta i antisosiale handlinger.

Jevnaldrende og skole

Antisosiale barn har vanskelig for å forholde seg prososialt til andre barn og tilpasse seg nye forventninger til atferd og arbeidsinnsats når de begynner i barnehagen eller på skolen (Dishion & Patterson, 2006). Gjennom sin atferd sender de kraftige sosiale signaler til omgivelsene slik at de ifølge Patterson (1982) reprogrammerer miljøet i barnehagen eller skolen til å bli konfliktfylt og avvisende. Ifølge Patterson fører den antisosiale atferden til at barn avvises av prososiale barn, samt til dårlige skoleprestasjoner og lav selvoppfatning. Det er den antisosiale atferden som skaper skolefaglige problemer, ikke omvendt (Patterson mfl. 1993). I 12–13-årsalderen etablerer antisosiale barn grupper som kjennetegnes av negative holdninger til skole og voksne autoriteter (Dishion & Patterson, 2006). Kontakten med slike grupper fører til ytterligere oppøving av eller 'trening' i antisosial atferd og øker også sannsynligheten for at de begynner å bruke rusmidler, skulke og begå kriminelle handlinger (Dishion mfl., 1999; Dodge mfl., 2006). Patterson og kollegers teoretiske modell for utvikling

av antisosial atferd danner grunnlag for behandlingsprogrammet 'Parent Management Training' (Patterson mfl., 1992; Patterson & Forgatch, 2000, 2005).

PMTO - behandlingsprogrammet

Parent Management Training (Oregon modellen: PMT-O) er et standardisert og strukturert behandlingsprogram for foreldre som har barn med atferdsvansker (Reid mfl., 2002)

Programmet har blitt oversatt og bearbeidet i en norsk behandlingsprotokoll (Solholm mfl., 2005) og kvalifiserte PMT-O terapeuter formidler, innøver og bekrefter alternative foreldreferdigheter slik at foreldre lærere å utnytte sine sterke sider og ta i bruk programmets kjernekomponenter.

Kjernekomponentene i programmet er foreldreferdigheter som gjennom forskning har vist seg å forebygge og motvirke antisosial atferd: 1) positiv involvering, 2) oppmuntring av ferdigheter, 3) problemløsning, 4) tilsyn og 5) grensesetting. Positiv involvering innebærer at foreldrene er oppmerksomme og viser interesse for å gi barnet omsorg (Forgatch & DeGarmo, 1999), oppmuntring av ferdigheter fremmer barnets kompetanse gjennom bruk av positive og betingede konsekvenser (Forgatch & DeGarmo, 1999) og problemløsning bidrar til at familiemedlemmene forhandler seg fram til enighet, etablerer husregler og avtaler konsekvenser for å følge eller bryte regler. Tilsyn beskytter barnet mot å delta i risikobetonete aktiviteter og handler om hvordan foreldrene holder seg orienterte om hva barnet holder på med (Patterson & Forgatch, 2005) og grensesetting formidler klare grenser for akseptabel atferd og motvirker avvikende atferd gjennom betinget bruk av moderate sanksjoner (Patterson, 1986).

I den første behandlingsfasen legger terapeuten vekt på å formidle støtte og forståelse for familiens situasjon og å etablere en samarbeidsrelasjon med foreldrene. Det legges vekt på å utvikle familiens sterke sider gjennom positive tilbakemeldinger og oppmuntrende foreldreferdigheter introduseres før fokuset rettes mot temaer som tilsyn, grensesetting og bruk av negative konsekvenser. Behandlingssamtalene er strukturerte, der tiden fordeles mellom å drøfte erfaringene med bruk av foreldreverktøyene og de innspill eller problemer som familien selv er opptatt av samt praktisk innøving av kjernekomponentene som de ikke behersker. I tillegg til å bygge på teori, er PMTO utviklet på grunnlag av omfattende forskning om utvikling av antisosial atferd og om effektene av behandling (Reid mfl., 2002). Et viktig ledd i behandlingen er monitorering og evaluering av integritet i betydningen lojal og kompetent praktisering av PMTO. Behandlingsintegritet blir målt som kompetent formidling av programmets kjernekomponenter i samsvar med kriteriene i "Fidelity of Implementation"

System” (FIMP, Knutson, Forgatch, & Rains, 2003). Video-opptak som består av ti minutters utdrag av definerte sesjoner blir evaluert i forhold til følgende fem dimensjoner: PMTO kunnskap, stukturering av timen, formidlingspraksis, prosessferdigheter og helhetlig kvalitet.

Forskning

Patterson mfl. (1993) oppsummerte at Oregon modellen hadde vært mest vellykket når det gjaldt å redusere ulydighet (trass) og andre antisosiale symptomer hos barn i alderen 3 til 8 år. Selv om behandlingen også hadde vist seg virksom i arbeidet med antisosial atferd hos ungdom (Bank, Marlowe, Reid, Patterson & Weinrott, 1991) så var resultatene for barn i alderen 9-12 år og ungdom i puberteten mindre oppmuntrende. Og selv om Oregon modellen av PMT også har vist seg å kunne bidra til bedre skoletilpasning, så har resultatstudiene i følge Patterson mfl. (1993) understreket behovet for å utvikle behandlingsmodellen i forhold til å bearbeide barnas skolefaglige problemer. Senere oppsummeringer av modellen bekrefter disse tendensene og utfordringene (Dishion & Patterson, 2006; Patterson, 2005). Det bør også nevnes at prinsippene fra PMTO har blitt anvendt som viktige ingredienser i behandlingsprogram som Multidimensional Treatment Foster Care (MTFC, Chamberlain, 1998).

I forskningslitteraturen er PMTO bedre etablert som forebyggende program enn som behandlingprogram, og metoden omtales ofte som teoretisk velfundert og empirisk validert. I oversiktsartikler omtaler Brestan & Eyberg (1998) og Eyberg mfl., (2008) PMT basert på Patterson, Reid, Jones & Congers (1975) modell som ”well established treatment” og underbygger dette med to undersøkelser av eldre dato (Bernal, Klinnert & Schultz, 1980; Patterson, Chamberlain & Reid, 1982). En nyere randomisert behandlingseffektstudie viste at PMTO førte til en signifikant reduksjon i foreldrenes vurderinger av barnas problematferd, men lærervurderingene ikke viste noen endring (Patterson & Forgatch, 1995). I et forebyggende prosjekt rettet mot barn av *separerte mødre* viste en randomisert undersøkelse at PMTO førte til positive observerte atferdsendringer (Martinez og Forgatch, 2001) til bedre skoleprestasjoner (Forgatch & DeGarmo, 2002), redusert problematferd over tid (DeGarmo, Patterson, & Forgatch, 2004) og reduksjon i norm- og regelbrytende atferd etter tre år (DeGarmo & Forgatch, 2005) og ni år (Forgatch, Patterson, DeGarmo & Beldavs, 2009). I et annet forebyggende prosjekt med fokus på barn i *stefamilier* viste det seg at foreldretrening etter PMTO prinsippene førte til redusert problematferd både hjemme og på skolen (Forgatch, DeGarmo, & Beldavs, 2005). Studiene bekreftet Pattersons (1982, 2002) teoretiske modell

som forklarer at foreldreferdigheter er virksomme endringsmekanismer, særlig grensesetting (eng. discipline) og oppfølging med milde negative konsekvenser og positiv involvering (eng. positive parenting).

På 1990 tallet viste også andre resultater fra kontrollerte evalueringsstudier at en kunne oppnå betydelig reduksjon i barns antisosiale atferd gjennom foreldreprogram og etterundersøkelser viste positive oppfølgingsresultater over perioder fra ett til syv år (Kazdin, 1997, Miller og Prinz, 1990, Patterson mfl., 1993, Forehand og Long, 1991, Webster-Stratton & Herbert, 1994). Det siste tiåret har foreldreprogrammene vært tema for flere systematiske forskningsoversikter og meta-analyser. Disse omtales i neste avsnitt.

Meta-analyser og systematiske kunnskapsoversikter

Lundahl, Risser og Lovejoy (2006) identifiserte i en meta-analyse 69 atferdsorienterte programmer som ble evaluerte med kontrollgruppedesign i perioden 1974-2003.

Resultatvariablene for barna omfattet økt lydighet eller samarbeid og redusert problematferd under tilsyn av foreldrene, mens det for foreldrene dreide seg om endringer i oppdragelsesferdigheter. Resultatvariablene omfattet også endringer i foreldrenes stress, kompetanse, forståelse av barneoppdragelse og generell tilpasning. Effektstørrelsen ved avslutning av terapiene var moderate, med Cohens $d = .42$ for barnas atferd og $d = .45$ for foreldrenes atferd. I gjennomsnitt deltok foreldrene i 9.77 (SD = 3.81) sesjoner som tilsvarte 14.74 (SD = 8.16) timers direkte terapeutkontakt, men verken antall sesjoner eller timer korrelerte med resultatvariablene (altså ingen dose-response sammenheng). Oppfølgingsdata ble samlet inn i 56 studier med kontrollgrupper etter gjennomsnittlig 7.72 (SD = 8.30) måneder. Resultatene viste at den gjennomsnittlige effektstørrelsen var lavere enn ved avslutningen av terapiene med Cohens $d = .21$ for barnevariablene og $d = .25$ for foreldrevariablene.

Moderator analysene i de atferdsorienterte programmene viste at barn og foreldre fra sosialt vanskeligstilte familier hadde mindre utbytte og endret seg i mindre grad enn de som var sosialt og økonomisk bedre stilt. Men barna i disse familiene klarte seg bedre hvis foreldrene ble behandlet individuelt ($d = .69$) heller enn å delta i foreldregrupper ($d = .34$). Barn i grupper med en relativ stor andel eneforsørgere endret seg mindre enn barn i grupper med forholdsvis færre eneforsørgere. Barn med klinisk signifikant forhøyet problematferd endret seg mer enn barn med moderate problemer eller barn i blandede grupper. Noe uventet hadde ikke alder noen signifikant modererende effekt på barnas utbytte av behandlingen, selv om effektstørrelsene var større for yngre enn for eldre barn (effektstørrelse for førskolealder d

= .44, for barnetrinnet $d = .31$ og mellomtrinnet $d = .27$). Lundahl mfl. (2006) konkluderte med at foreldreprogrammer som retter seg mot barns eksternaliserte atferdsproblemer var en relativt robust behandlingsform som ga moderate effektstørrelser ved avslutningen av terapiene, men små effektstørrelser i oppfølgingsperioder opp til ett år etter avsluttet behandling.

McCart, Priester, Davies & Azen (2006) sammenlignet atferdsorienterte foreldretreningsprogram (BPT; behavioral parent training) med kognitiv atferdsterapi (KBT) i 71 studier i perioden 1973-2004. Studien omfattet barn i alderen 6-12 år (BPT = 7 studier og KBT = 21 studier). Foreldrene var den viktigste informantgruppen for BPT mens lærere var hovedinformanter i KBT studiene. BPT programmene hadde en høyere vektet gjennomsnittlig effektstørrelse ved avslutningen av intervensjonene enn KBT ($d = 0.45$ vs. $d = 0.23$). Det var ingen positiv sammenheng mellom alder og effektstørrelse i BPT, men for KBT var det en slik sammenheng. Forfatterne konkluderte med at BPT var mest effektiv for de yngste barna fordi de var mer avhengig av sine foreldres veiledning og støtte mens KBT var mer virksomt enn BPT for eldre. Det kan forklares med at barn etter hvert som de blir eldre utvikler abstrakte, kognitive ferdigheter og tilbringer mer tid sammen med og påvirkes mer av venner (McCart mfl., 2006; Weisz, 1997). Et problem med å sammenligne BPT med KBT var imidlertid at de to tilnærmingene ble evaluert av forskjellige informantgrupper (foreldre vs. lærere) og aldersspennet var forholdsvis stort sammenlignet med andre meta-analyser.

Reyno & McGrath (2006). I denne meta-analysen identifiserte Reyno og McGrath (2006) hvilke variabler som i 31 behandlingsstudier (publisert 1980-2004) predikerte kliniske behandlingsresultater hos barn med eksternaliserende problematferd. Det var bare lav familieinntekt som ga store standardiserte effektstørrelser, men det var også utslag for barnas atferdsmessige belastningsgrad ved inntak og mødrenes psykopatologi. Forfatterne oppsummerte at resultatene av foreldretrening ofte ble påvirket av variabler som bare indirekte involverte barnet. Det viste seg at sosio-økonomisk status inkludert familiens inntekt var den viktigste enkeltprediktoren for behandlingsutfall mens effektstørrelsen for mødrenes mentale helse var liten. Mødres depresjon så ut til å påvirke kvaliteten av mor-barn interaksjonen, mødrenes oppfatning av barnets atferd og måten de satte grenser på. Overdreven hardhendt oppdragelse viste seg å være den sterkeste prediktoren for barnas eksternaliserende problematferd (Reyno & McGrath, 2006).

Piquero, Farrington, Welsh, Tremblay og Jennings (2008). Piquero mfl. (2008) gjennomførte en meta-analyse av 55 publiserte og upubliserte evalueringer (1976-2008) av foreldretreningsprogrammer, men også 'home-visiting' programmer som forebygger atferdsproblemer hos barn (Olds, Robinson, O'Brian et al., 2002). Inklusjonskravet var at undersøkelsene skulle ha et eksperimentelt design og begrenset seg til familier med barn under 5 år eller undersøkelser der gjennomsnittsalderen for barna var cirka 5 år. Gjennomsnittlige effektstørrelser for hver studie ble beregnet basert på resultatene ved avslutningen av behandlingene (heller enn oppfølgingsresultater). Best representert blant foreldreprogrammene var The Incredible Years Program (Webster-Stratton & Hammond, 1997), The Triple P Positive Parenting Program (Sanders, 1999) og the Parent-Child Interaction Therapy (PCIT, Brestan, Eyberg, Boggs & Algina, 1997; Brinkmeyer & Eyberg, 2003). Den systematiske forskningsoversikten viste at tidlig familie/ foreldretrening var en virksom intervensjon i arbeidet med å redusere atferdsproblemer blant små barn. Den uvektede gjennomsnittlige effektstørrelsen var $d = 0.46$ og dersom en lot store undersøkelser telle mer enn små, så var den vektete effektstørrelsen $d = 0.45$. Det var ikke forskjeller i behandlingsutfall mellom foreldretrenings- og 'home visiting' programmer

Forelderapporteringer viste de mest positive resultatene ($d = .34$) fulgt av observasjonsdata ($d = .28$) og lærervurderinger ($d = .24$) og ellers gikk tendensen i retning av at de eldste studiene hadde høyere effektstørrelser enn de nyere. De eneste signifikante moderatorene for effektstørrelse var publikasjonsland og utvalgsstørrelse. Undersøkelser gjennomført i USA hadde en marginalt større signifikant effektstørrelse enn undersøkelser gjennomført i andre land. Utvalgsstørrelse var den sterkeste prediktoren for variasjoner i effektstørrelser, og undersøkelser som ble gjennomført med færre enn 100 deltakere hadde signifikant større effektstørrelser ($d = .47$) enn de med flere enn 100 deltakere ($d = .21$).

Heterogene resultatvariabler var et problem i denne som i andre meta-analyser av foreldreprogram. De teoretiske og operasjonelle definisjonene av aggresjon, antisosial atferd og norm- og regelbrytende atferd (eng. delinquency) varierer så mye at det ikke er klart om det er de samme resultatvariablene som ble målt i de ulike studiene og over tid. Meta-analyser kan også vekte resultatvariabler forskjellig, slik at for eksempel lærervurderinger betraktes som mer uavhengige i en familieintervensjon enn foreldrenes og at systematiske "unbiased" observasjoner igjen kan være mer nøyaktige enn lærervurderinger. Noen meta-analyser løser dette problemet gjennom å beregne gjennomsnittlige effektstørrelser av alle resultatvariablene (Farrington & Welsh, 2003).

Kaminski, Valle, Filene og Boyle (2008). Kaminski mfl. (2008) gjennomførte en komponent-analyse der de identifiserte virksomme program- eller behandlingskomponenter i 77 publiserte evalueringer av foreldreprogram for å forbedre atferd og tilpasning hos barn i alderen 0-7 år. Dette var den første meta-analysen som sammenlignet spesifikke programkomponenter, for eksempel forskjeller i innhold eller formidlingsform. For å kunne trekke representative konklusjoner ble også studier med ikke-randomiserte forskningsdesign inkludert i analysen. Resultatene omfattet endringer umiddelbart etter avsluttet behandling og viste en gjennomsnittlig effektstørrelsen på $d = 0.34$. Den samlede effektstørrelsen for foreldrevariabler ($d = .43$, kunnskaper, holdninger, verdier, self-efficacy og atferd) var større enn effektstørrelsen for barnevariablene ($d = 0.30$, eksternalisering og internalisering, skolefaglig eller sosialt kompetent atferd). Størst var effektstørrelsen for internaliserende problematferd og minst for sosiale ferdigheter.

Undersøkelser med randomiserte design rapporterte signifikant større effektstørrelser enn andre design og undersøkelser der en sammenlignet gruppene ved oppstart hadde mindre effektstørrelser enn de som ikke målte forskjeller ved basislinjen. Frittstående foreldreprogram hadde større effektstørrelser enn de som var del av en intervensjonspakke. To metodologiske faktorer ga utslag: undersøkelser der kontrollgruppen ikke fikk noe tilbud og undersøkelser der foreldre var eneste informant; begge var forbundet med høyere effektstørrelser. Videre viste det seg at behandlinger som forutsatte at foreldrene skulle praktisere nye ferdigheter med sitt eget barn i sesjonene rapporterte signifikant høyere effektstørrelser enn programmer der barna ikke deltok i timene.

For gruppen av barn med eksternalisert problematferd (men ikke for barn med andre problemkjennetegn) var det signifikante effektstørrelser for komponenten 'pausetid' (timeout) og konsekvente responser på barnets atferd. I behandlingen av eksternalisert problematferd var det også virksomt å lære foreldrene hvordan de kunne samhandle positivt med sine barn når det ikke var konflikter, for eksempel hvordan de kunne vise entusiasme og positiv oppmerksomhet for ønsket atferd, hvordan samhandle på barnets nivå under lek, la barnet lede i lek og la barnet velge fritidsaktiviteter. Konsekvente reaksjoner handlet om betingede reaksjoner, dvs. å reagere (forutsigbart) på en bestemt problematferd hver gang den forekom og med den samme konsekvensen. Som i andre meta-analyser (Lundahl mfl., 2006) viste det seg at å tilby foreldrene tilleggstenester ga mindre effekter, muligens fordi det trakk foreldrenes oppmerksomhet bort fra selv å lære seg nye oppdragsferdigheter. Det viste seg

også at standardiserte modellprogrammer ikke ga høyere effektstørrelser enn generell behandling uten slike kjennetegn.

For barna var effektstørrelsene høyere for internalisert enn for eksternalisert problematferd. Disse effektstørrelsene var høyere enn for kognitive og skolefaglige ferdigheter, som igjen var høyere enn effektstørrelsene for sosiale ferdigheter og pro-sosiale atferd. Forklaringene på lavere effektstørrelser for sosiale ferdigheter kan være at barna i denne aldersgruppen (0-7 år) hadde begrensede muligheter for å praktisere ferdighetene med venner og andre personer utenom familien. Bare én av de fire mest robuste prediktorene handlet om grensesetting. To andre, positiv samhandling og emosjonelle kommunikasjonsferdigheter (bl.a. aktiv lytting, reflektere hva barnet sa og hjelpe barnet med å identifisere og mestre følelser) handlet om hvordan en kunne forbedre kvaliteten av foreldre-barn relasjonene gjennom å øke foreldrenes aktivitet og positivitet i forhold til barnet, eventuelt redusere foreldrenes negativitet (for eks. kritikk og sarkasmer). Læring i kontekst var effektivt, det vil si at foreldre og barn fikk anledning til å praktisere nye ferdigheter sammen i behandlings-timene. Forfatterne hevdet at det var viktig å forbedre eksisterende programmer ved å legge til komponenter som hadde vist seg virksomme på tvers av programmer heller enn å skifte ut hele programmet. Komponentanalyser gjør det også mulig å eliminere komponenter som viser seg å være lite virksomme. Meta-analysen konkluderte med at foreldrebehandling endrer foreldreatferd og forebygger eller redusere barns problematferd. De gjennomsnittlige effektstørrelsene var større for foreldreferdigheter enn for barnas atferd, og effektstørrelsene var mindre for foreldrenes atferd og ferdigheter enn for kunnskap, holdninger eller 'self-efficacy'.

Meta-analysene var ikke uten videre sammenlignbare fordi de varierte mye med hensyn til inklusjonskriterier for programmene, hvilket tidsrommet de dekket, barnas alder, og hva de sammenlignet. Likevel er det mulig å peke på noen felles resultater og konklusjoner. Selv om meta-analysene ikke sammenlignet spesifikke programmer så konkluderte de med at gruppen atferdsorienterte foreldreprogrammer hadde en positiv behandlingseffekt på barn med alvorlige atferdsproblemer og deres foreldre. Resultatene pekte i retning av at effektstørrelsene var moderate ved avslutningen av terapiene og små i oppfølgingsstudiene, og de var større for foreldrevariabler enn for variabler som målte endringer i barnas atferd. Det var en tendens i retning av at de yngre barna endret seg mer enn de eldre, og barn med stor problembelastning ved inntak endret seg mer enn barn med færre problemer. Det var generaliseringseffekter av behandlingen til barnas internaliserte problematferd og andre

funksjonsområder, mens spørsmålet om generalisering av atferdsendring fra hjem til skole, ikke kunne besvares entydig. Det var indikasjoner på at barn av eneforsørgere, barn av foreldre med lav sosio-økonomisk status og barn av foreldre med dårlig psykisk helse (særlig deprimerte mødre) kom dårligere ut enn barn av foreldre med få sosiale, økonomiske og personlige problemer. Videre ga enkeltprogrammer bedre resultater enn de som inngikk i større programpakker og familier med sosiale problemer hadde bedre utbytte hvis de ble behandlet individuelt heller enn at de deltok i foreldregrupper. Tilleggskomponenter som individualbehandling av barn og tilleggstjenester til foreldrene økte ikke behandlingsutbyttet for barna. Enkelte av analysene viste også at undersøkelsen design og kontekst kunne påvirke resultatene slik at undersøkelser som bare baserte seg på foreldrevurderinger, eldre undersøkelser, undersøkelser med små utvalg, undersøkelser der kontrollgruppen ikke fikk noe behandlingstilbud og undersøkelser gjennomført i USA, ga bedre resultater enn de øvrige undersøkelsene. Komponentanalysen underbygger dessuten funn fra effektstudier som viste at positive og proaktive forelderferdigheter hadde en gunstig effekt på barna og at betingede (eller konsekvente) reaksjoner og 'pausetid' hadde spesifikk positiv effekt på utagerende barn.

Multisystemisk terapi (MST)

Multisystemisk terapi (MST) er en familie- og nærmiljøbasert og retter seg mot alvorlige atferdsproblemer som lovbrudd og rusmisbruk hos ungdom i alderen 13-17 år. Hensikten er å forebygge videre kriminell aktivitet og andre former for antisosial atferd samt å forebygge at ungdom fengsles eller plasseres utenfor hjemmet. Multisystemisk terapi (MST) kan beskrives som et behandlingsprogram, men omfatter også en nøye spesifisert organisasjonsstruktur og et faglig nettverk av MST-team og enheter (Christensen & Mauseith, 2007; Henggeler mfl., 1998, 2009). I Norge er teamene organisatorisk knyttet til det den statlige barnevernstjenesten mens henvisningene kommer fra den kommunale barnevernstjenesten. MST kan rubriseres som et hjemmebasert tjenestetilbud som bygger på prinsippet om "familiebevaring" (eng. family preservation).

Hver behandlingsenhet (eng. treatment site) består av ett eller flere team som ledes av en klinisk veileder (eng. clinical supervisor) og som består av 3-4 terapeuter, som hver har ansvaret for å følge opp 3- 6 familier. Teamet og/eller familiens terapeut er tilgjengelig for familien 24 timer i døgnet, 7 dager i uken. I perioder holdes det daglig kontakt gjennom hjemmebesøk eller over telefon.

Det teoretiske grunnlaget for MST

MST har et multiteoretisk grunnlag som beskrives av Henggeler og medarbeidere (1998, 2009) på tre nivåer. På det øverste nivået finnes Bronfenbrenners (1979) sosialøkologi og teoretiske modeller som bygger på empiriske studier av utviklingen av antisosial atferd (bl.a. Elliott, Huizinga & Ageton, 1985). Det neste nivået består av teorier som mer avgrenset er knyttet til de enkelte tiltakskomponentene, for eksempel systemisk familieteori og kognitiv atferdsmodifikasjon (Haley, 1976; Minuchin, 1974). På det tredje nivået er det formulert ni teoretiske behandlingsprinsipper som styrer valg og prioritering av tiltak, og disse brukes for å fastslå om intervensjonene samsvarer med de teoretiske forutsetningene for MST.

Det sosialøkologiske grunnlaget for MST kan beskrives som en progressiv gjensidig tilpasning mellom ungdom i utvikling og et foranderlig oppvekstmiljø (Bronfenbrenner, 1979). Det gjensidige påvirkningsforholdet mellom ungdommene og deres umiddelbare miljø (omtalet som mikro- og mesosystemet) preges igjen av den større sosiale konteksten som interaksjonen foregår i (omtalt som ekso- og makrosystemet). I MST anvendes det økologiske perspektivet både på ungdoms utvikling og for å planlegge helhetlige og sammenhengende tiltak og tjenester.

Atferdsproblemer defineres i MST som problematiske transaksjoner i og mellom sosiale systemer der ungdommen deltar sammen med familien, medelever og lærere, jevnaldringsgruppen og andre. Det sosialøkologiske grunnlaget for MST markeres også gjennom prinsippene om at (Henggeler mfl., 1998): (1) ungdommens atferd kan bare forstås innenfor helheten som består av deres nærmiljø eller kontekst, (2) individet har kapasitet til å forandre seg, (3) individets fenomenologi, det vil si dets subjektive definisjoner og fortolkning av egen situasjon er viktig for å forstå utvikling og atferd og (4) den beste måten å teste hypoteser på er gjennom naturalistiske eksperimenter. Hypoteser om hvilke tiltak som kan endre ungdommens atferd og situasjon kan med andre ord testes gjennom å endre den sosiale konteksten (jf Bronfenbrenner, 1979). En konsekvens av det utviklingsøkologiske perspektivet er prinsippet om ”familiebevaring” der tjenestene blir tilpasset familiemedlemmenes behov og den enkelte families verdier, tro og kultur.

En annen teoretisk hovedpill er årsaksmodellstudier som forklarer hvordan antisosial atferd utvikles hos barn og unge (Elliott mfl., 1985). Blant de viktigste risikofaktorene for utvikling av kriminalitet (Henggeler, 1991) og rusmiddelmisbruk (Henggeler, 1997) er: forbindelser med antisosiale jevnaldrende, dyfunksjonelle familierelasjoner, skoleproblemer

og manglende støtte fra naboer og lokalmiljøet (Elliott mfl. 1985). Ungdommenes forhold til venner og skolen er avgjørende for utviklingsforløpet. Avvisning av prososiale jevnaldrende og kontakt med avvikende venner er en sterk prediktor for dårlig psykososial tilpasning mens skolen er en viktig arena for utvikling av sosial og skolefaglige kompetanse. I MST har imidlertid skolerettede tiltak lavere prioritet enn de familierettede (Rodick & Henggeler, 1980). MST-terapeuten skal prioritere å utvikle familiens kapasitet for problemløsning slik at de i samarbeid med skolen kan løse ungdommens skoleproblemer (med mindre problemene er akutte og alvorlige eller skolen ber terapeuten om assistanse).

De overordnede teoretiske perspektivene suppleres med teorier som mer avgrenset er knyttet til de ulike tiltakskomponentene som for eksempel systemisk familieteori (Haley, 1976, Minuchin, 1974), atferdsøkologi (Munger, 1993), sosial interaksjons læringsteori (Patterson, 1979), økosystemisk intervensjonsteori (Cooper, Smith & Upton, 1994) og kognitiv atferdsmodifikasjon (Beck, 1995; Meichenbaum, 1993). Disse komponentene har sitt eget teorigrunnlag samtidig som de kan innpasses i en sosialøkologisk referanseramme. Det er imidlertid en forutsetning at de ikke bryter med prinsippet om at systemiske intervensjoner skal prioriteres framfor individuelt rettede behandlingstiltak. Familierapien kritiseres imidlertid for ikke å ta tilstrekkelig hensyn til hvordan atferdsproblemene påvirkes av individuelle kjennetegn eller kontekstuelle forhold utenfor familien. (Henggeler mfl., 1998).

Teoretiske behandlingsprinsipper

På det tredje teoretiske nivået integreres teori og forskning i ni behandlingsprinsipper. Innenfor rammene av disse teoretiske prinsippene prioriteres evidensbaserte metoder. Empirisk validerte metoder som "best practices" trekkes inn etter behov, og kan erstattes hvis forskning viser at andre metoder er mer effektive. De ni behandlingsprinsippene er: 1) Den primære hensikten med utredningen er å forstå den gjensidige tilpasningen mellom de identifiserte problemene og deres bredere systemiske kontekst, dvs. vurdere hvordan identifiserte problemer opprettholdes av familiens nåværende sosiale miljø, 2) terapeutiske kontakter vektlegger det positive og bruker systemisk styrke for å fremme forandring i kontakten med familiene, 3) intervensjoner skal fremme ansvarlig atferd og redusere uansvarlig atferd hos familiemedlemmene, 4) intervensjonene er "her og nå" og handlingsfokuserte, rettet mot konkrete og veldefinerte problemer, 5) intervensjoner rettes mot atferdssekvenser innenfor eller mellom multiple systemer, 6) intervensjonene tilpasses utviklingstrinnet og ungdommens utviklingsmessige behov, 7) intervensjonene bør kreve daglig eller ukentlig innsats av familiemedlemmene, 8) intervensjonenes effektivitet evalueres

kontinuerlig fra flere perspektiver og 9) intervensjonene bør utformes slik at de fremmer generalisering og langtids vedlikehold av terapeutisk forandring. Behandlingsprinsippene (som er nærmere beskrevet i appendiks 1) definerer et relativt stort handlingsrom og åpner for kreativitet i valg av tiltak. Som en prinsippstyrt behandlingsmodell legger MST opp til individuelt tilpassede mål og tiltaksplaner som er forenlig med de overordnede teoretiske prinsippene.

Multisystemisk terapi

MST er et intensivt, målrettet og tidsavgrenset behandlingstilbud med følgende kjennetegn: 1) arbeidet foregår i familiens naturlige omgivelser – hjemme, på skolen og i nabolaget, 2) behandlingen er tidsavgrenset, og varer vanligvis mellom 3 og 5 måneder pr. familie, 3) det arbeides vanligvis i team med 3-4 terapeuter, som hver har ansvar for å følge opp 3 til 6 familier hver og 4) terapeuten eller teamet er tilgjengelig 24 timer i døgnet, 7 dager i uka og det holdes daglig kontakt med familiene, ansikt til ansikt eller over telefon. MST forutsetter et forpliktende samarbeid med minst én av ungdommens foresatte og god kontakt med lokale barne- og ungdomstjenester. Systemiske intervensjoner prioriteres de en tar hensyn til hvordan alle deltakerne i ungdommens økosystem bidrar til å opprettholde problemet og analyserer hvilken funksjonell sammenheng det er mellom deres atferd og ungdommens identifisert problemer. Terapeutene og de foresatte klargjør hva ungdommen oppnår med sin atferd og sammen søker de etter alternative og mer akseptable måter å nå målene på. Individuelle behandlingstiltak som kognitiv atferdsmodifikasjon kan være aktuelle når de systemiske intervensjoner ikke gir ønsket resultat.

Det forventes ikke at alle ungdommens og familiens problemer skal være løst når behandlingen avsluttes, men at familien har økt sin kompetanse slik at de kan fortsette å løse problemer på egen hånd. Filosofien i MST går ut på at den mest effektive og etiske måten å hjelpe ungdom på, er gjennom å hjelpe deres familier, og at dette lykkes bedre når en støtter heller enn klandrer dem. Familien skal settes i stand til å løse sine egne problemer og foreldrene får rollen som langsiktige forandringsagenter. Samtidig tar en høyde for at det kan være konflikter mellom ungdom og foreldre, og at situasjonen kan oppleves forskjellig av ungdom, foreldre og andre sentrale personer i det sosiale nettverket. Å gi opp familier eller stemple dem som ”motvillige” eller ”umotiverte” blir ansett for å være et lite aktuelt alternativ og det gjøres mye for å støtte foreldrene slik at de står løpet ut. Arbeidet med å utvikle, styrke og frigjøre familiens ressurser, er sentralt. MST skal formidle ”hjelp til selvhjelp” og MST-terapeutene skal derfor ikke gjøre noe som familiemedlemmene selv er i stand til å mestre.

Innsatsen i familien skal imidlertid komme ungdommen tilgode og derfor er endringer i hans eller hennes atferd den viktigste resultatindikatoren.

I MST legges det stor vekt på å monitorere i hvor stor grad terapeutene følger de ni behandlingsprinsippene og dette måles gjennom bruk av Treatment Adherence Measure scales (TAM, Henggeler & Borduin, 1992). En person som ikke deltar i terapiene (for eksempel en kontoransatt) ringer opp familiene én gang i måneden og ber dem svare på 26 spørsmål om innholdet i møtene mellom familiene og MST terapeuten. Informasjonen gjør det mulig å vurdere i hvilken grad arbeidet med den enkelte familie samsvarer med den idealtypiske MST-modellen og tilbakemeldingene fra familiene blir brukt i veiledning av terapeutene for å måle graden av behandlingsintegritet og forebygge avdrift fra programmet. (Henggeler mfl., 1998, 2009).

Forskning

I en av de tidlige effektstudiene av MST (Henggeler, Melton & Smith, 1992) ble ungdom som hadde begått alvorlig ungdomskriminalitet fordelt tilfeldig slik at én gruppe fikk MST mens en annen fikk "vanlige tjenester" fra "Youth Services". Gjennomsnittsalderen var 15,2 år og 77% var gutter. Da ungdommene ble fulgt opp etter 60 uker, viste det seg at de som hadde fått MST i signifikant mindre grad var blitt arrestert, de var i mindre grad plassert utenfor hjemmet og de rapporterte om mindre kriminalitet. Foreldrene i MST gruppen rapporterte også om bedre familiesamhold og at ungdommen var mindre aggressiv i forhold til jevnaldrende. MST var effektiv selv om ungdommene varierte med hensyn til etnisk bakgrunn, alder, sosial klasse og kjønn. I en ny oppfølging etter 2 år (Henggeler, Melton, Smith, Schoenwald, & Hanley, 1993) viste det seg at ungdommene i MST gruppa også denne gangen var arrestert signifikant færre ganger enn ungdommene i sammenligningsgruppen.

I en senere undersøkelse (Borduin mfl., 1995) ble resultatene fra MST sammenlignet med resultatene fra individualterapi med ungdom som hadde begått alvorlige lovbrudd. Gjennomsnittsalderen var 14,8 år og 67,5% var gutter. MST hadde en signifikant gunstigere effekt på relasjoner og samhandling i familien enn individualterapi. Det ble også registrert en reduksjon i foreldrenes rapportering av egne psykiske problemer (symptomatologi) og av ungdommenes atferdsproblemer. Resultatene etter 4 år viste at MST også var mer effektiv i arbeidet med å forebygge framtidige lovbrudd, inkludert voldskriminalitet. MST viste seg å være like effektiv for jenter som for gutter og for ungdom med forskjellig etnisk bakgrunn. En senere undersøkelse viste at etter 13.7 år hadde MST ungdommene signifikant mindre

tilbakefall til kriminalitet enn sammenligningsgruppen som fikk individualterapi (Schaeffer & Borduin, 2005)

Øvrige evalueringsresultater viser redusert internalisert og eksternalisert problematferd hos ungdom (Borduin mfl., 1995; Henggeler mfl., 1997), reduksjon i rusmisbruk (Henggeler, Pickrel & Brondino, 1999) reduksjon av psykiske symptomer hos foreldre (Borduin mfl., 1995; Scherer, Brondino, Henggeler, Melton & Hanley, 1994), økt sosial støtte, samhold og tilpasning samt færre konflikter og redusert fiendtlighet i familien (Borduin mfl., 1995). Resultatstudiene fra USA viste en nedgang i re-arrestasjoner blant ungdom som hadde fått MST på mellom 25 og 70 prosent (Henggeler mfl., 1992, 1997).

Mens MST studier i Norge (Ogden & Halliday-Boykins, 2004; Ogden & Amlund-Hagen, 2006) og New Zealand (Curtis, Ronan, Heiblum, & Kylie, 2009) viste positive resultater, så viste en replikasjonsstudie i Sverige at MST verken ga bedre eller dårligere resultater enn andre intervensjoner til ungdom med alvorlige atferdsvansker (Sundell mfl., 2006, 2008). Ungdommene i den svenske evalueringsstudien viste en generell reduksjon i atferdsproblemer på tvers av behandlingsbetingelsene etter syv måneder, men det var ingen signifikante gruppeforskjeller. I en oppfølgingsstudie viste det seg at resultatene etter 7 måneder viste samme tendens to år etter inntak. Löfholm, Olsson, Sundell og Hansson (2009) konkluderte derfor med at MST ikke hadde noen mer positiv langtidseffekt enn andre tiltak og tjenester som ble tilbudt målgruppen i Sverige.

Systematiske kunnskapsoversikter – meta-analyser av MST

I dette avsnittet presenteres meta-analyser av MST. Det er brukt forholdsvis mye plass på denne gjennomgangen fordi resultatene har vært motsetningsfylte, og særlig én av analysene har utløst en lengre debatt om innhold og fortolkning av analysen.

Curtis, Ronan og Borduin (2004) utførte en meta-analyse av forskningen om MST der inklusjonskriteriene var at (1) studiene var publiserte i utvalgte forskningstidsskrifter, (2) MST ble praktisert i overensstemmelse med MST prinsippene, (3) randomisering av deltakere på MST og en kontrollgruppe, (4) et klinisk utvalg med ungdommer som hadde antisosiale atferdsproblemer og/eller psykiatriske symptomer, (5) registrering av data før og etter behandling, eventuelt med oppfølgingsdata og (6) bruk av tester som kunne anvendes i en meta-analyse (middel-verdier, standardavvik og/eller Fisher's F-ratio). Forfatterne identifiserte syv studier med 708 deltakere og 35 MST terapeuter i fagfellevurderte artikler fra

1987-2002. Deltakerne var i hovedsak gutter (70%) som bodde sammen med minst én av sine biologiske foreldre og de fleste hadde lav sosio-økonomisk status. Blant deltakerne i studiene var 54% afro-amerikanske. Blant MST ungdommene fullførte 75-100% behandlingen mens den tilsvarende andelen var 56-100% for kontrollgruppene.

Effektstørrelsene ble beregnet som et gjennomsnitt av alle Cohen's d verdier i alle studiene og varierte mellom $d = .40$ og $d = .70$ med et gjennomsnitt på $d = .55$. De fant større effektstørrelser i effektstudier ($d = .81$) enn i effektivitetsstudier i ordinær praksis ($d = .26$). Mål på familierelasjoner hadde de største effektstørrelsene ($d = .57$), fulgt av effektstørrelser på individuelle mål ($d = .28$, for eksempel YSR og CBCL) mens effektstørrelsene for vennerelasjoner ikke var signifikante ($d = .11$). Dette kan forklares med at MST i motsetning til individuell behandling av ungdom primært retter seg mot familieprosesser slik at forandringene først kan registreres der og at disse igjen predikerer forandringer i ungdommenes atferd (Huey mfl., 2000; Mann, Borduin, Henggeler & Blaske, 1990). Blant modererende variabler framhevet de behandlingsintegritet (Henggeler mfl., 1997) mens mediatorer omfattet forbedring av vennerelasjoner (Huey mfl., 2000) og familiens engasjement (Schaeffer & Borduin, 2005). MST har også blitt testet på andre målgrupper med blandede resultater, og forfatterne konkluderte med at det er behov for mer empirisk støtte før en eventuelt kan konkludere at programmet er virksomt i behandling av ungdom med rusproblemer og i familie- og nærmiljøbehandling som alternativ til innleggelse av ungdom med akutte psykiatriske problemer. Blant begrensningene ved meta-analysen nevnte Curtis mfl. (2004) at det var et lite antall studier i analysen og at resultatene derfor måtte betraktes som tentative. I framtidige studier anbefalte forfatterne å legge mer vekt på å måle graden av kontakt med prososiale venner, kognitive prosesser som attribusjonsstil og attribusjonsskjevhet samt skoleprestasjoner og deltakelse i organiserte fritidsaktiviteter.

I en kommentar til meta-analysen (Curtis mfl., 2004) skrev Henggeler (2004) at det var åpenbart at effektstørrelsene var vesentlig mindre i vanlig praksis som ble gjennomført i lokalmiljøer med praktikere som fikk minimal direkte veiledning fra programutviklerne enn i effektstudier som var universitetsbaserte med doktorgrads studenter som terapeuter og med kontinuerlig tilbakemelding og veiledning fra MST utviklerne. Målsettingen var imidlertid at det skulle være mulig å implementere og evaluere MST uten programutviklernes medvirkning.

Resultatene fra meta-analysen til Curtis mfl. (2004) ble utfordret i en protokoll i Campbell Collaboration og Cochrane library av Littell, Popa og Forsythe (2005). Inklusjonskravet var at undersøkelsene skulle ha benyttet et kvantitativt gruppedesign med

randomisert fordeling av deltakerne på tiltaksalternativene (RCT), samt anvendt standardiserte undersøkelsesinstrumenter, slik at pre-post resultater skal kunne sammenlignes på tvers av undersøkelsene. Utvelgelseskriteriene medførte at 8 studier ble inkludert, 6 fra USA, én fra Canada og én fra Norge. Hensikten med den systematiske forskningsoversikten var å sette fokus på det de oppfattet som et misforhold mellom forskning og praksis i arbeidet med meta analyser. Forskningsoversikten kom til et annet resultat enn primærstudiene som meta-analysen bygde på og konkluderte med at det var prematurt å trekke konklusjoner om MSTs effektivitet sammenlignet med andre tjenester. Littell m.fl. (2005) stilte ikke krav til at studiene i analysen skulle ha vært publisert i tidsskrifter med fagfellevurderinger noe som medførte at den upubliserte canadiske MST studien (Cunningham, 2002; Lescheid & Cunningham, 2002) ble inkludert. Littell (2005a) påpekte at det kan ha fått betydning for at hennes konklusjoner avviker fra de som er trukket i andre meta analyser som Curtis mfl. (2004). Mens syv av de åtte primærstudiene konkluderte med at det var signifikante forskjeller i favør av MST på én eller flere resultatvariabler så fant Littell m.fl. (2005) at selv om dataene hadde en tendens til å favorisere MST, så var ikke forskjellene statistisk signifikante. I meta analysen re-analyserte Littell og kolleger resultatene variabel for variabel i 21 enkeltanalyser. I hver av analysene inngikk maksimalt fem undersøkelser, og i de fleste tilfellene var det bare tre eller to. Med et så lite antall primærstudier kreves det stor homogenitet i design og resultater for å kunne påvise statistisk signifikante forskjeller¹. Men i dette begrensede utvalget hadde studiene høy heterogenitet, det vil si at de skilte seg fra hverandre på områder som undersøkelsesgruppe, observasjonsperiode, implementering og design. Med bare åtte undersøkelser var det vanskelig å avgjøre om MST var signifikant bedre enn tilbudet i sammenligningsgruppen. Littell m.fl. (2005) kommenterte dette slik : "Since statistical power is low, we cannot conclude that MST is not more effective than other services" (s.22)². Den canadiske MST studien (Cunningham, 2002; Lescheid & Cunningham, 2002) veide tungt i denne meta-analysen fordi den hadde mange deltakere (N=409) og utgjorde 21-45% av undersøkelsespopulasjonen i meta analysene. Samtidig rapporterer den klart mer negative resultater sammenlignet med de andre evalueringene. Den canadiske studien skilte seg også fra andre MST studier ved at den omfattet ungdom med færre problemer, tiltakene i kontrollgruppen hadde en høyere kvalitet enn i USA og implementeringskvaliteten var lavere (Lescheid & Cunningham, 2002).

¹ Ifølge Lipsey (2000) har testen (Q-test) som brukes i slike analyser svak 'power' og en trenger minst 20 separate studier for å kunne si noe om gjennomsnittresultater.

Littell m.fl. (2005) påpekte det de oppfattet som forskningsmetodiske svakheter ved MST studiene som 1) inkonsistent rapportering av antall saker i ulike publikasjoner fra samme prosjekt, 2) inkonsekvent parvis kobling mellom MST og sammenligningssaker, 3) uklare randomiseringsprosedyrer, 4) ustandardiserte observasjonsperioder i oppfølgingsstudiene, og 5) subjektiv definisjon av behandlingsavslutning. Littell m.fl. hevdet blant annet at selektivt frafall (dropout) i MST studiene hadde ført til at resultatene var blitt mer positive enn dersom dette hadde vært håndtert på andre måter. Henggeler, Schoenwald, Swenson og Borduin (2006) hevdet i et tilsvarende at det er viktig å skille mellom ulike faser av programutviklingen når en skal vurdere behandlingseffektivitet, og at en må ta hensyn til variasjoner i behandlingsintegritet mellom utprøvningssteder. Resultatene fra MST studier viste ifølge Henggeler mfl. (2006) at det første året var særlig krevende for MST når det ble implementert på nye steder og uten programutviklernes direkte medvirkning. Evalueringer av programmer rett etter oppstart kan derfor føre til at en undervurderer dets potensiale. Tilsvaret fra Littell (2005b) kretser mye rundt den canadiske studien der hun i motsetning til Henggeler mfl. (2006) hevdet at data ikke ga grunnlag for å konkludere med at programmets "modenhet" påvirket resultatene i positiv retning.

Littell m.fl. (2005) skilte ikke mellom effekt og effektivitetsstudier i sin meta-analyse og mente at dette skillet var uklart i MST forskningen. Mens Henggeler (2004) hevdet at behandlingsintegritet i MST i flere undersøkelser hadde vist seg å korrelere positivt med behandlingsutfall så hevdet Littell mfl. (2005) at TAM ikke diskriminerte mellom MST og andre intervensjoner. Det som ble målt med TAM var i følge Littell mfl. ikke unikt for MST, slik som for eksempel engasjement, behandlingsdeltakelse og terapeutisk allianse. Videre hevdet de at data ikke støttet hypotesen om at MST var mer effektivt på noen behandlingssteder enn på andre og viste til at i to av tre undersøkelser med flere utprøvningssteder, ble det ikke rapportert resultater fra det enkelte utprøvningssted. De mente også at det var galt å påberope forskjeller mellom behandlingssteder i ettertid (post hoc) og etterlyste flernivåstudier ('multilevel studies') som kan redegjøre for forskjeller i behandlingseffekt mellom terapeuter og behandlingssteder i den samme undersøkelsen (såkalt nesting effects). Littell m.fl. (2005) hevdet også at resultater fra det enkelte utprøvningssted ('site effects') var uinteressante i en meta-analyse der en aggregerte resultatene fra all tilgjengelig forskning.

I en kommentar til meta analysen hevdet Ogden og Amlund-Hagen (2006a) at Littell og kolleger (2005) stilte urimelig høye standardiseringskrav til design, variabler og prosedyrer i undersøkelser som skal inngå i en meta analyse. De påpekte også at måten data ble re-analysert på, kan ha påvirket konklusjonene slik at de stemte dårlig med primærstudiene som

inngikk i analysen. Med så få og heterogene studier i meta analysen mente Ogden og Amlund-Hagen (2006) med støtte i rådene fra The Cochrane Collaboration at det kanskje var prematurt å foreta en meta-analyse av MST. Og hvis MST forskningen faktisk var så mangelfull og med den kvalitetssvikt som Littell og kolleger (2005) beskrev, undergravde de samtidig sine egne konklusjoner fordi en meta analyse blir aldri bedre enn de data, metoder, og studier den bygger på.

I en senere meta-analyse med et litt annet utgangspunkt enn de som er nevnt ovenfor undersøkte Drake, Aos og Miller (2009) hvilke tiltak som kunne redusere tilbakefall til kriminalitet og hvilken kostnad-nytte effekt de kunne ha for staten Washington. Analysen var en videreføring av en tidligere meta-analyse av blant annet evidensbaserte program for reduksjon av kriminalitet (Aos, Miller & Drake, 2006). Meta-analysen inkluderte alle engelskspråklige evalueringer etter 1970 uavhengig av om de var publisert eller ikke. Evalueringene måtte ha en kontrollgruppe, og selv om randomiserte studier ble prioritert, så omfattet meta-analysen også utvalgte ikke randomiserte studier. Det var imidlertid en forutsetning at en kunne analysere ut fra intention-to-treat (ITT) designet, dvs. alle som var fordelt på henholdsvis intervensjon og kontrollgruppe forble i betingelsen uavhengig av deltakelse (eventuelt med behandlingsfravall mindre enn 10%). I studier med oppfølgingsdata inkluderte forfatterne effektstørrelser for den lengste oppfølgingsperioden og data der en korrigerte for relevante kovariater. Det ble korrigert for kvaliteten på forskningsdesignet slik at høyest rangering ble gitt til studier av godt implementerte program med randomisert fordeling på intervensjons- og kontrollgruppe. I analysene foretok de også en vekting, der effektivitetsstudier av programmer som ble gjennomført i ordinær praksis veide mer enn effektstudier. En lang rekke tiltak og programmer ble omfattet av analysen, inkludert MST. Analysen av MST var basert på 10 studier og viste en reduksjon i kriminalitet på 7,7%. Dette var svakere enn resultatene som ble oppnådd med Multidimensional Treatment Foster Care (MTFC) og Funksjonell familieterapi (FFT) der den beregnede reduksjonen i kriminalitet var på henholdsvis 17, 9% og 18,1%.

Westen mfl. (2004, 2005) som kritiserte de fleste evidensbaserte programmer for å være korttidsbaserte vedgikk samtidig at forskere kan teste lengre behandlingsforløp i randomiserte studier. I den sammenhengen viste de til Henggeler og kollegers forskning samtidig som de hevdet at evidensbasert behandling bare unntaksvis brukte så få eksklusjonskriterier og behandlet så intensivt som MST. Westen mfl. (ibid.) omtalte også Henggeler og kollegers MST som et kreativt eksempel på hvordan en kan utlede behandling fra grunnforskning så vel som fra klinisk erfaring, og å integrere begreper og teknikker på

tvers av teorier med fokus på klienter som hører til de som er vanskeligst å behandle i sitt lokalmiljø. Noen av funnene i MST fant Westen mfl. (2005) inspirerende, blant annet den betydelige reduksjonen i tilbakefall til kriminalitet, og særlig voldskriminalitet etter 4 år (Borduin mfl., 1995) og etter 13.7 år (Schaeffer & Borduin, 2005).

Samtidig påpekte Westen mfl. (2005) problemene med å overføre MST til ordinær praksis. Sammenlignet med effektstudiene viste effektivitetsstudiene av MST bare en tredjedel av effektstørrelsen, noe som kunne skyldes at effektstudiene ble gjennomført med klinikere som var doktorgrads studenter og som arbeidet i et forskningsmiljø med en saksmengde på bare to til tre familier og under grundig veiledning (Curtis mfl., 2004). En styrke ved MST studiene var også at behandlingen ble sammenlignet med genuin alternativ praksis (TAU, *treatment-as-usual*). Dermed støttet resultatene forskernes generaliseringer om behandlingens relative effekt vis-a-vis vanlig praksis. Men Westen mfl. (ibid.) påpekte samtidig at resultatene på flere variabler var langt fra optimale som når det var relativt små effektstørrelser på målinger av individuell psykopatologi sammenlignet med familievariablene og fortsatt høy forekomst av psykopatologi. Westen mfl. (2005) kritiserte programutviklerne for ikke å ha evaluert om resultatene endret seg hvis en forlenget behandlingen utover 3-4 måneder og for at de ikke hadde evaluert hvilke komponenter i MST som var effektive, uviktige eller uheldige. Dette framstår som relevant kritikk og konstruktive innspill til videre MST forskning.

Oppsummert har MST vist seg å ha langsiktig reduserende effekt på rus- og voldsrelatert kriminalitet og andre lovbrudd, men også på antall arrestasjoner, fengslinger og plasseringer utenfor hjemmet. Behandlingen har også vist seg å føre til forbedringer av foreldre-ungdom og øvrige familierelasjoner, til redusert aggresjon overfor venner og redusert kontakt med avvikende venner. Langtidseffektene av MST er hovedsaklig basert på registerdata som fanger opp arrestasjoner og dommer for lovbrudd. De mest positive evalueringsresultatene er oppnådd av programutviklerne og deres kolleger, mens uavhengige replikasjonsstudier i andre land viser både positive (Ogden & Amlund-Hagen, 2006; Curtis mfl., 2009) og nøytrale resultater (Sundell mfl., 2008; Löfholm mfl., 2009). Spesifikke meta-analyser av MST er problematiske på grunn av få og heterogene studier, men det finnes én som konkluderte med at MST var virksomt i arbeidet med å redusere emosjonelle og atferdsmessige problemer hos ungdom og familiemedlemmene (Curtis mfl., 2004) og én som konkluderte med at MST ikke kom bedre ut enn behandlingene og tiltakene den ble sammenlignet med (Littell mfl., 2005).

Oppsummering og diskusjon

Parent Management Training Oregon modellen har et solid teoretisk og empirisk fundament i Pattersons (1982, 2002) sosial interaksjons læringsteori (SIL modellen) og i forskningen ved Oregon Social Learning Center. I følge PMTO er manglende eller dysfunksjonelle foreldreferdigheter den viktigste drivkraften bak utviklingen av antisosial atferd og omgang med antisosiale venner kan føre til en ytterligere opptrapping. PMTO behandlingen legger stor vekt på å endre barnas atferd gjennom at foreldrene systematisk lærer å bruke fem grupper av oppdragelsesferdigheter.

PMTO inngår i familien av atferdsorienterte foreldreprogram som kan vise til robuste, men moderate resultater, særlig over tid. Ved siden av å ha inspirert beslektede behandlingsprogrammer har PMTO kunnet vise til solide forskningsresultater i forebyggende prosjekter både på kort og lang sikt. Komponentanalysen bekreftet studier som viser at foreldrebehandling har størst endringspotensiale i forhold til foreldrenes atferd, noe mindre i forhold til barnas internaliserte og eksternaliserte problematferd og minst i forhold til endringer i barnas skolefaglig funksjonsnivå og sosial kompetanse. Analysene tyder også på at grensesetting (eng. discipline) i betydningen 'pausetid' (eng. timeout) og betingede konsekvente reaksjoner er viktige for å redusere utagerende atferd sammen med økt innsats i forhold til positiv involvering og proaktive kommunikasjonsferdigheter i konfliktfrie situasjoner.

MST har et multi-teoretisk grunnlag og kan omfatte kognitiv-atferdsmessige tilnærminger, atferdsmodifikasjon, foreldretrening og pragmatisk familieterapi og andre intervensjoner som har en rimelig evidensbasis (Henggeler & Lee, 2003). MST har en betydelig fleksibilitet når det gjelder innholdskomponenter og derfor er behandlingen operasjonalisert som 'adherence' til ni teoretiske behandlingsprinsipper som regulerer behandlingsplanleggingen. Et sentralt prinsipp er imidlertid at intervensjonene skal være systemiske, altså primært rette seg mot ungdommenes kontekst der flere personer kan være endringsagenter så som foreldre, søsken, familie, venner, lærere og andre. MST som behandlingsmodell analyserer og bearbeider ungdommenes problemer i forhold til den sosiale konteksten i og utenfor familien. Evalueringene viser at MST påvirker ungdommenes atferd gjennom å endre samhandling og samhold i familien og gjennom å regulere kontakten med jevnaldrende, mens forholdet til skolen er mindre utforsket.

MST synes også å være en robust intervensjon med få eksklusjonskriterier og som fungerer i forhold til de sammensatte problemene som ungdommene henvises for. Resultatene synes å variere med kontrollbetingelsene, der kvaliteten og intensiteten av ordinære tjenester som en sammenligner med påvirker den relative effektstørrelsen av MST. En viktig utfordring handler om hvordan en skal få MST behandlingen implementert og virksom i ”det virkelige liv” og uten programutviklernes direkte medvirkning. Det kan virke som om MST er avhengig av et solid støtteapparat som ivaretar kvalitetssikring av opplæring, veiledning og behandlingsintegritet. Vellykket behandling er også avhengig av et integrerende lokalmiljø der en lykkes med å mobilisere MST-vennlige tjenester og en inkluderende skole eller arbeidsliv. Det handler også om å kunne kontrollere påvirkningene fra antisosiale subkulturer eller begrense lokalmiljøets utstøttingsmekanismer når det kommer ønsker eller krav om at ungdommene må interneres eller omplasseres.

Meta-analysene av MST utløste like mye en diskusjon om hva meta-analyser bør inneholde og hvordan de bør gjennomføres som en diskusjon om Multisystemisk Terapi virker eller ikke. Et viktig spørsmål var om meta-analyser kan komme til et helt annet resultat enn primærstudiene som de bygger på. Diskusjonen illustrerer at det ikke finnes allment aksepterte kriterier for hvilke kvalitetskrav som skal stilles til meta-analyser og til de randomiserte kontrollerte studier som de bygger på. Foreløpig er det ikke gjennomført komponentanalyser av MST, og det hadde vært nyttig med analyser som viste hvilken effekt de ulike ingrediensene i behandling har på resultatene eller om en forlengelse av behandlingstiden kunne gi større effekt.

Selv om ikke meta-analyser er tema for denne avhandlingen, så er de nyttige fordi de oppsummerer og formidler gjennomsnittlige eller standardiserte behandlingsresultater på tvers av mange primærstudier. De kan være et viktig korrektiv til enkeltstudier, eller de kan bekrefte funn og dermed underbygge konklusjoner i primærstudier. Meta-analyser kan også gi nyttig informasjon om hvilke tilnærminger og programmer som gir de mest lovende resultatene, eller formidle hvilke komponenter eller ingredienser som synes å være mest virksomme. Videre gir meta-analyser muligheter for å analysere sterke og svake sider ved ulike typer forskningsdesign, variabelvalg og sammensetning av utvalg. Bedre enn enkeltstudier gir de også muligheter for å analysere hvilke forhold som modererer behandlingsutbyttet. På denne måten kan resultater fra meta-analyser representerer en nyttig referanseramme for nye primærstudier, og gi innspill til viktige forskningshypoteser. Samtidig er det viktig å være oppmerksom på meta-analysenes begrensninger. En av svakhetene går ut

på at en slår sammen gamle og nye studier, der de gamle ofte har mindre utvalg, svakere design og flere feilkilder. Hvis en slår sammen gode og dårlige studier, så kan de beste studiene få sin motvekt i dårlige studier, noe som kan føre til at gjennomsnittresultatene får liten informasjonsverdi. Kazdin (2008) understreker at en i diskusjonen om evidensbaserte metoder og program må skille mellom ulike klientgrupper, problemområder og behandlingsmåter. Hvis en som i visse meta-analyser og systematiske kunnskapsoversikter slår sammen og diskuterer ulike klientgrupper, problemområder og behandlingsmåter under ett, kan en risikere å trekke konklusjoner som ikke er valide for visse problemområder (for eksempel alvorlige atferdsproblemer), visse målgrupper (for eksempel barn og unge) og visse program (som kan variere mye med hensyn til hvordan behandlingsprotokoller eller manualer er utarbeidet).

6. Presentasjon av undersøkelsene

Fem publiserte studier utgjør grunnlaget for denne avhandlingen. Den første studien er en randomisert evalueringsstudie av Oregon modellen for Parent Management Training (PMTO) der de kliniske resultatene ble evaluert ved avslutningen av behandlingene (Ogden & Amlund-Hagen, 2008). De tre neste studiene er knyttet til Multisystemisk terapi (MST) og omfatter en pre-post analyse av kliniske resultater ved avslutning av terapiene (Ogden & Halliday-Boykins, 2004) og en oppfølgingsstudie to år etter inntak til behandling (Ogden & Amlund-Hagen, 2006). Den tredje studien tar for seg kjønnsforskjeller i inntakskjennetegn, komorbiditet og behandlingsutfall i MST (Ogden & Amlund-Hagen, 2009). Den fjerde og siste studien er en retrospektiv kassustudie beskriver implementeringen av PMTO og MST i Norge (Ogden, Amlund-Hagen, Askeland & Christensen, 2009).

Studie nr. 1. Evaluering av Parent Management Training – Oregon modellen

Formål

Hensikten med undersøkelsen var å sammenligne resultatene fra behandlingsprogrammet PMTO med ordinære tjenester i et geografisk bredt sammensatt utvalg av norske barn med alvorlige atferdsproblemer og deres familier. Programmets effektivitet ble evaluert i ordinær praksis i Barnevernstjenesten og psykisk helsevern (Barne- og ungdomspsykiatriske poliklinikker).

Metode

Design og deltakere

Deltakere i studien var 112 barn i alderen 4 til 12 år med en gjennomsnittsalder på 8.44 år (SD = 2.13 år) og deres foreldre. De aller fleste var av norsk opprinnelse og 90 var gutter (80,4%) mens 22 var jenter (19,6%). Barna ble tilfeldig fordelt med 59 på PMTO og 53 på andre behandlingstiltak. I studien bidro 33 PMTO terapeuter og 51 terapeuter som behandlet barna i sammenlignings- eller kontroll gruppen. Kontrollgruppen fikk tiltakene som ordinært ble tilbudt målgruppen i de to tjenestene før PMTO ble introdusert blant annet familie-terapi (n = 10), Marte-meo (n = 6), atferdsterapi (n = 2), humanistisk-eksistensiell terapi (n = 1) og annet (n = 27) (informasjon manglet for 6 familier).

Inntak til behandlingen i begge grupper var basert på kliniske vurderinger heller enn diagnoser og inklusjonskriteriene omfattet atferd som var nevnt i DSM-IV manualen (APA, 1994) som opposisjonell trasslidelse eller atferdsforstyrrelse eller en problembeskrivelse i tråd med disse, så som aggresjon, norm- og regelbrudd eller problematferd i skolen. Det var ingen familier som ble ekskludert og ingen som avsto å delta i behandlingen noe som kan henge sammen med at det var få eksklusjonskriterier og at foreldrene på forhånd hadde fått god informasjon om hvem behandlingstilbudet var beregnet på. De standardiserte T-skårene på atferdsvurderinger viste at barna skåret høyt over gjennomsnittet blant norske barn med en CBCL (Child Behavior Checklist) skåre på 66, 17 (SD = 10,0) og TRF (Teacher Report Form) på 68, 51 (SD = 9, 78). Blant de 104 familiene det forelå opplysninger om, var 62 foreldre gift eller samboende mens 42 var eneforsørgere.

Prosedyre

Rekrutteringen til studien foregikk i perioden januar 2001 til april 2005 og behandlingen ble gjennomført i perioden oktober 2001 til mai 2006. Foreldrene samtykket skriftlig til å delta i studien og ga også samtykke til at barnas lærere eller førskolelærere kunne bidra med sine vurderinger. PMTO terapeutene ble rekrutterte gjennom en åpen invitasjon til de som var ferdig utdannet, og rekrutterte en kontrollterapeut fra egen eller en annen organisasjon.

Familiene gjennomgikk kartleggingsprosedyren ved inntak og ble deretter randomisert til intervensjonsgruppen (PMTO) eller sammenligningsgruppen. Randomiseringen ble foretatt på hvert behandlingssted og av terapeutene selv. De siste tolv familiene som var i inntaksfasen da rekrutteringen ble stoppet ble likevel inkludert og fordelt uten å vente på den andre familien i paret. Behandlingen ble avsluttet når familien og terapeuten var enige om at målene for behandlingen var nådd, at barnets problematferd var under rimelig kontroll, eller hvis familien ga uttrykk for at den ønsket å avslutte. Kartleggingen ved avslutningen av terapiene var identisk med inntakskartleggingen.

Kartleggingsinstrumenter

Noen av kartleggingsinstrumentene var felles for evalueringsstudiene av PMTO og MST. Det gjaldt Achenbach's (1991) Child Behavior Checklist (CBCL), Teacher Report Form (TRF) og familiens evaluering av tilbudet (Family Satisfaction Scale, Lubrecht, 1992). Instrumentene som var felles i de to studiene presenteres nedenfor og og deretter beskrives instrumentene som bare ble benyttet i PMTO eller MST studien.

Foreldres atferdsvurderinger. "The Child Behavior Checklist" ble fylt ut av ungdommenes foresatte (CBCL, Achenbach, 1991) og leddene vurderes på en 3 punkts Likert skala (0 = sjelden/aldri, 1 = noen ganger, 2 = ofte/alltid) der høye skårer indikerer høy problembelastning. Skalaen har blitt oversatt, standardisert og validert i norske undersøkelser (Nøvik, 1999; Ogden & Amlund-Hagen, 2006). I begge MST studiene viste CBCL gode psykometriske egenskaper med α ' verdier mellom .81 og .94. I PMTO studien ble delskalaene for internalisert (α = .89 ved både pre og post) eksternalisert (α = .93 ved både pre og post) problematferd benyttet sammen med total problemskåre (α = .94 ved pre og .96 ved post).

Læreres atferdsvurderinger. "The Teacher Report Form (TRF, Achenbach, 1991c) er ekvivalent til CBCL, og har blitt oversatt og validert i andre norske undersøkelser (Lurie, 2006; Ogden & Amlund-Hagen, 2006). I PMTO studien var det en tilfredsstillende indre konsistens målt med alpha verdier for eksternalisert (α = .95 ved pre og post), internalisert (α = .85 ved pre og .88 ved post) problematferd, mens for totalskåren var α = .92 ved pre og .93 ved post.

Familiens evaluering av kvaliteten av tilbudet De foresatte besvarte 12 spørsmål om hvor fornøyde de var med behandlingen de hadde mottatt (Family satisfaction scale, Lubrecht, 1992). Vurderingene ble foretatt langs en 4 punkts Likert skala og en høy skåre signaliserte høy tilfredshet (α = .83 i MST studien mens i PMTO studien var α = .87 for den første (N = 75) og α = .87 (N = 30) for den andre av foreldrene.

Øvrige instrumenter i PMTO studien

Social Skills Rating Scale (SSRS) (Gresham & Elliott, 1990) er en vurderingsskala for foreldre og lærere som fokuserer på sosiale ferdigheter og skolefaglig kompetanse (bare lærere). Versjonen for barn (barnehage til 6. trinn) ble brukt i denne studien og skalaen har vist tilfredsstillende reliabilitet og validitet i tidligere norske studier (Ogden, 2003; Sørli, Amlund-Hagen & Ogden, 2008). Foreldreversjonen har 38 ledd (α = .86 og .83 ved pre og post) mens lærerversjonen har 30 ledd (α = .85 og .82 ved pre og post).

Foreldrenes daglige rapportering (Parent Daily Report, PDR) Over tre påfølgende dager registrerer foreldrene barnas atferdsproblemer før og etter behandlingen (Patterson mfl. 1982). Den primære omsorgsgiveren ble ringt opp og svarte 'ja' eller 'nei' på om noen av de 34 hendelsene i indeksen hadde forekommet i løpet av de siste 24 timene. En sumskåre ble laget for de 3 dagene samlet og anvendt i analysene. Cronbach's alpha var .81 før behandlingen og .84 ved endt behandling.

Observasjon og koding av strukturerte samspillsoppgaver i familien. Strukturerte samspillsoppgaver ble videofilmet i 30 minutter for barn over 8 år og i 25 minutter for barn under 8 år. Begge grupper fikk en problemløsningsoppgave og en evalueringsoppgave der de drøftet hvor bra de hadde klart oppgavene. Familiene til de yngste barna deltok også i oppgaver med frilek, opprydning og venting. Familier med eldre barn brukte mer tid på problemløsningsoppgaven og skulle i tillegg planlegge en hyggelig aktivitet den neste uken. Video-opptakene ble kodet av reliable kodere som ikke visste hvilken gruppe familiene tilhørte. Foreldre-barn interaksjonen ble kodet etter "the Family Peer Process Code" (Stubbs, Crosby, Forgatch & Capaldi, 1998) der en numerisk registrerer interaksjoner mellom familiemedlemmene ut fra varighet og sekvens. Det ble registrert hvor ofte barnet initierte en sekvens eller kjede av negative interaksjoner. Sekvensen besto av et negativt verbalt eller non-verbalt budskap som ble besvart med en negativ reaksjon fra den foresatte innen 6 sekunder. Observasjonsdataene fra video-opptakene ble skåret på to måter, som mikrososiale registreringer (FPPC) og som globale vurderinger foretatt av kodere (koders inntrykk skala).

Koders inntrykk skala er et spørreskjema der koderne gir en global vurdering av foreldrenes oppdragelsesferdigheter i de strukturerte samspillsituasjonene. Skjemaet ble fylt ut av koderne etter at de hadde fullført den mikrososiale kodingen av video-opptakene. Skjemaet har ulike versjoner for barn under og over 8 år, men i begge blir fem foreldreferdigheter vurdert sammen med barns lydighet (eng. compliance). Foreldreferdigheten 'oppmuntring av ferdigheter' ble ikke benyttet på grunn av dårlige psykometriske egenskaper i denne studien. Før den samlede effekten ble analysert, ble skårene standardisert for hver av versjonene (over og under 8 år).

Behandlingsintegritet. Behandlingsintegritet ble målt som kompetent gjennomføring av behandlingen i samsvar med kriteriene i "Fidelity of Implementation System" (Knutson, Forgatch, & Rains, 2003, jf. kapittel 5). FIMP kodere var norske PMTO spesialister som hadde fått opplæring slik at de skåret reliabelt i forhold til kodere i USA. Sesjonene som ble valgt var introduksjon og problemløsning knyttet til grensesetting og oppmuntring av ferdigheter. Skårene ble beregnet som et gjennomsnitt av de fire sesjonene. I undersøkelsen ble 41% av filmene kodet og cirka 20% av videoopptakene ble kontrollert for interrater reliabilitet. Tilfredsstillende reliabilitet mellom kodere i Norge og USA ble sikret gjennom månedlige kalibrering og reliabilitetskontroll gjennom videokonferanser. De norske terapiene ble transkribert og oversatt til engelsk før de ble evaluert av de amerikanske koderne.

Intervensjon

PMTO behandlingen fulgte prinsippene i behandlingsprotokoller fra Oregon Social Learning Center (Forgatch, 1994) og disse ble oversatt til norsk og supplert med norsk materiale (Askeland, Christiansen & Solholm, 2005). Selv om behandlingen fulgte PMTO protokollen så var behandlingen fleksibel og tok hensyn til familienes individuelle behov. Hver behandlingstime skulle også fokusere på ett av fem forhåndsdefinerte kjernekomponenter som omfattet (Solholm m.fl., 2005): 1) positiv involvering og kommunikasjon, 2) oppmuntring og positive konsekvenser, 3) problemløsning, 4) grensesetting og 5) tilsyn (monitoring). Foreldrene fikk muligheter til å lære og praktisere ferdighetene gjennom rollespill og øvelser, og når de mestret ferdighetene fikk de i oppgave å prøve ut dem ut i hjemmesituasjonen. Det ble avsatt tid til å bearbeide foreldrenes vansker, blant annet for å kunne øke deres overskudd og motivasjon for å hjelpe barnet. PMTO terapeutene som bidro i studien hadde gjennomgått en 18 måneders opplæring, fullført fem terapier og hadde blitt sertifisert på grunnlag av ekspertvurderinger av video-opptak av utvalgte behandlingssesjoner. Terapeutene fikk også regelmessig veiledning og deltok jevnlig i vedlikeholdsseminarer i prosjektperioden.

Statistiske analyser.

De primære resultatvariablene i studien var eksternalisert problematferd, sosial kompetanse, lydighet og 'negative interaksjoner initiert av barna' sammen med foreldreferdighetene grensesetting, positiv involvering, problemløsning og monitorering (tilsyn). De sekundære resultatvariablene var internalisert problematferd og total problemskåre på CBCL og TRF, skolefaglig prestasjoner og behandlingsintegritet målt med FIMP. Analysene var i henhold til 'intention-to-treat' slik at de inkluderte alle saker som opprinnelig ble randomisert. I en serie kovarians analyser ble behandlingseffektivitet testet med dosering, inntaksskårer, aldersnivå og kjønn som kovariater. Videre ble en stimodell med 'maximum likelihood estimation' brukt for å teste både spesifikke indirekte og individuelle effekter. I modellen predikerte behandlingsbetingelse effektiv disiplin og denne variabelen predikerte igjen tre resultatvariabler; barnas lydighet, frekvensen av negative interaksjoner initiert av barna og en samlevariabel for eksternalisert problematferd. Samlevariabelen (eng. externalizing composite) inkluderte z-skårer for de eksternaliserte subskalaene fra CBCL og TRF samt summen av PDR skårer over tre dager. En høy skåre på samlevariabelen indikerte at både foreldre og lærere vurderte at barnet hadde høy eksternalisert problematferd.

Missing data og frafall

Blant de 112 familiene som ble rekruttert til studien av slo 7 PMTO familier og 8 familier i sammenligningsgruppen å delta i kartleggingen ved avslutningen av terapiene. Familiene som ikke deltok på den avsluttende kartleggingen skilte seg ikke signifikant de som medvirket. 'Expectation maximization' prosedyren ble benyttet for å predikere totalskåren hvis det forelå mangelfulle data på en variabel. Hvis en 'missing-completely-at random' analyse ikke lyktes eller hvis løsningen ikke konvergente, ble regresjon benyttet for å predikere verdier for beregning av totalskåren. Hvis data manglet fullstendig på en variabel for en respondent, ble saken vurdert som 'completely missing'.

Resultater

Sammenligning av grupper ved inntak. Blant de 112 familiene som deltok i studien bidro 97 familier (87%) ved avslutningen av behandlingene. Den gjennomsnittlige behandlingstiden var 47 uker. Ved inntak (basislinje) var barna i PMTO gruppen signifikant eldre enn barna i sammenligningsgruppen (9.05 vs. 7.77 år). Videre var det flere eneforsørgere i sammenligningsgruppen (24/47) enn i PMTO gruppen (18/57). Det var ingen andre forskjeller mellom gruppene på øvrige variabler av betydning for undersøkelsen.

Dosering. Fordi PMTO familiene fikk signifikant flere behandlingstimer enn kontrollfamiliene (27,67 vs. 20,41) og barna i sammenligningsgruppen fikk flere timer behandling enn PMTO barna, så ble doseringen (antall timer i behandling) brukt som kovariat i analyser av behandlingseffektivitet. Dosering korrelerte ikke signifikant med CBCL eller TRF total problemskåre og predikerte derfor ikke behandlingsutfall i denne studien.

Behandlingsutfall. Foreldre som hadde mottatt PMTO rapporterte om mindre *eksternalisert* problematferd hos barna sine ved avslutningen av behandlingen og endringer i absolutte T-skårer fra pre til post var 6.57 for PMTO barna og 4.39 for barna i sammenligningsgruppen. Barna i PMTO gruppen skåret også signifikant lavere på *CBCL totalskåren* enn kontrollgruppen. Foreldrenes daglige rapportering av problematferd (PDR) over 3 dager viste marginalt signifikant lavere skåre for barna i PMTO gruppen og i direkte sammenligninger av aldersgrupper viste det seg at barna som var yngre enn åtte år hadde en signifikant reduksjon i PDR skårer, men det gjaldt ikke for barna som var eldre (simple effects analyse: når gruppene over og under åtte år ble sammenlignet hver for seg på tvers av behandlingsbetingelsene).

Lærerne rapporterte om signifikant høyere sosial kompetanse i PMTO gruppen enn i sammenligningsgruppen ved avslutningen av terapiene. Barna i PMTO gruppen som var yngre enn 8 hadde i følge sine lærere mindre *eksternalisert* problematferd og lavere totalskåre

på TRF mens det ikke var noen forskjell for de eldre barna (8-12 år) når aldersgruppene ble sammenlignet direkte (jf simple effects).

En marginalt signifikant interaksjonseffekt mellom behandling og alder ble også funnet for de yngste barna på variabelen barns 'lydighet' (eng. compliance) vurdert av observasjonskodere. Videre var det en interaksjonseffekt som viste at jentene i PMTO gruppen ble mer positivt vurderte enn jentene i sammenligningsgruppen (det var ingen alderseffekt på denne variabelen), men effekten var ikke signifikant når gutter og jenter ble sammenlignet hver for seg på tvers av behandlingsbetingelsene.

Foreldreferdigheter. Foreldrene i PMTO gruppen mestret i følge kodernes inntrykk grensesetting signifikant bedre enn foreldrene de ble sammenlignet med. Foreldrene til de yngste PMTO barna ble også vurdert mer positivt i forhold til tilsyn (monitoring) med barnet enn foreldrene til de yngste i sammenligningsgruppen. Interaksjonseffekt som ble avdekket for 'problemløsning' viste noe uventet at foreldrene til de eldste barna i sammenligningsgruppen skåret høyere etter endt behandling enn foreldrene i PMTO gruppen. Det var imidlertid ingen forskjell for de yngre barna, og det var heller ingen gruppeforskjeller i foreldreferdigheten 'positiv involvering'.

Familienes tilfredshet med behandlingen. Foreldrenes evalueringer av behandlingen gikk i PMTO familienes favør. Mens 77% av foreldrene i PMTO gruppen ville anbefale behandlingen til andre, så var det tilsvarende tallet for kontrollfamiliene 38%. Enveis variansanalyse viste at foreldrene (dvs. hovedrespondenten) i PMTO gruppen var signifikant mer tilfreds med behandlingen enn familiene i sammenligningsgruppen. Dette gjaldt uavhengig om det var én eller begge foresatte som deltok i behandlingen og vurderingen. Terapeutvurderingene viste imidlertid ingen gruppeforskjeller i familienes tilfredshet. Siden svarene fra familienes hovedrespondent korrelerte med skårene på behandlingsintegritet (FIMP) var det grunn til å anta at foreldrenes tilfredshet økte med stigende behandlingsintegritet. Høy behandlingsintegritet (FIMP skårer) var også korrelert med høye skårer på foreldreferdighetene positiv involvering og grensesetting

Analyse av en modell for indirekte effekt. I samsvar med Pattersons (1982) SIL modell testet vi hypotesen om at forbedring av foreldreferdigheter indirekte formidlet effekten av behandlingsbetingelsene på barnas atferd i en stimodell. De avhengige variablene i modellen var en samlevariabel for vurderinger av barnas eksternaliserte problematferd, frekvensen av negative interaksjoner initiert av barna i strukturerte samspillsituasjoner (kodet på grunnlag av observasjonsdata) og barnas lydighet (kodere globale vurderinger eller inntrykk). De uavhengige variablene var aldersnivå (over og under 8 år), strukturert samspillsbetingelse

(forskjellig for barn over og under 8 år), inntaksskårer for foreldrenes ferdigheter i grensesetting og behandlingsbetingelse. Stimodellen viste at kompetent grensesetting predikerte mer lydighet, færre negative interaksjoner initiert av barna og mindre eksternalisert problematferd. Men den eneste signifikante, *spesifikke* indirekte effekten av behandlingsbetingelsene via foreldrenes ferdigheter i grensesetting fant vi i forhold til barnas lydighet. Foreldrenes ferdigheter i grensesetting hadde størst sammenheng med barns lydighet, mindre med 'negative interaksjoner initiert av barna' og minst sammenheng med barnas samlede problematferd.

Praktisk/klinisk signifikans og effektstørrelser. Effektstørrelser (Cohens d) ble brukt som indikatorer på praktisk-klinisk signifikans. Effektstørrelsene var i hovedsak små til moderate. Effektstørrelsen var $d = 0.16$ for eksternalisering og $d = 0.20$ for total problemskåre på CBCL. Effektstørrelsene for lærerrapporterte sosiale ferdigheter (SSRS) var $d = 0.47$ og for foreldrenes ferdigheter i grensesetting $d = 0.30$. Den eneste effektstørrelsen om kunne karakteriseres som stor var for foreldrenes tilfredshet med behandlingen, som tilsvarte $d = 0.83$. Effektstørrelsene må imidlertid vurderes med en viss varsomhet fordi et 95% konfidensintervall for noen av variablene inkluderte 0.

Studie 2, 3 og 4: Evaluering av Multisystemisk Terapi (MST)

Målsettingen med evalueringen var for det første å undersøke om positive behandlingsresultater fra USA kunne replikeres i Norge for ungdom med alvorlige atferdsproblemer og for det andre å undersøke i hvilken grad MST på sikt ga resultater som var bedre enn de en oppnådde gjennom tiltakene som Barnevernstjenesten ellers iverksatte for denne målgruppen. De norske behandlingsresultatene ble evaluert ved avslutning av terapiene (eller etter cirka 6 mdr.) (Ogden & Halliday-Boykins, 2004) og i en oppfølgingsstudie to år etter inntak (Ogden & Amlund-Hagen, 2006). Siden det norske MST prosjektet hadde sin forankring i barnevernet, var forebygging av plasseringer utenfor hjemmet og reduksjon i ungdommenes problematferd hjemme, på skolen og i fritiden primære resultatvariabler. En tredje studie fokuserte på kjønnsforskjeller i henvisningsårsaker og bakgrunnskjennetegn ved inntak til MST (Ogden & Amlund-Hagen, 2009). Studien analyserte også kjønnsforskjeller i forekomst av tilleggsp problemer ved inntak, om jentene endret seg like mye som guttene i løpet av behandlingen og om foreldrene til gutter og jenter vurderte behandlingskvaliteten forskjellig.

Metode

MST - design og deltakere

Ett hundre ungdommer, deres foresatte, og lærere deltok i evalueringsstudien av MST. Familiene ble rekruttert fra de fire første fylkeskommunale utprøvningsstedene som tok imot tilbudet om å opprette MST-enheter i sine fylker. Det ble rekruttert 100 deltakere til studien og hver MST-enhet bidro med 25 saker. Av de 100 deltakerne, var 37 jenter og gjennomsnittsalderen var ved inntak 15 år. Praktisk talt alle ungdommene var født og oppvokst i Norge. Ved inntak til prosjektet bodde 46 av ungdommene i to-foreldre familier, 38 bodde sammen én forsørger mens 15 var plassert utenfor hjemmet. Blant de viktigste henvisningsårsakene var 'identifisert som alvorlig atferdsproblem' (64%), alvorlige regelbrudd som rømming og skulk (53%), rusmisbruk (50%), lovbrudd (37%), trusler om å skade seg selv eller andre (36%) og vold i hjemmet (29%). Mange av ungdommene ble henvist av flere årsaker.

Oppfølgingsstudien to år etter inntak omfattet 75 ungdommer fra utprøvningssteder. Av disse var 27 jenter og gjennomsnittsalderen 17.07 år (SD = 1.36). Frafallet i denne gruppen var svært lavt slik at 69 av de opprinnelige 75 familiene deltok etter to år.

I analysen av kjønnsforskjeller deltok 117 ungdommer ved to behandlingssteder hvorav 41 var jenter (35%) og 76 gutter (65%). Deltakergruppen besto av de 62 ungdommene som hadde blitt randomisert til MST i behandlingseffektstudien av MST (Ogden & Amlund-Hagen, 2006) og i tillegg 55 ungdommer som fikk MST ved to behandlingssteder det påfølgende året. Det eneste som skilte de 55 ungdommene som ble rekruttert uten å bli randomisert fra de 62 i effektstudien var at de var cirka 6 måneder yngre og de skåret lavere på selvrapportering av internalisert problematferd og oppmerksomhetsproblemer (YSR). Gjennomsnittsalderen var 14.7 år for jentene og 14.2 for guttene og det var ingen signifikante forskjeller i deres sosiale bakgrunnskjennetegn.

Kartleggingsinstrumenter

Som nevnt ble kartleggingsinstrumentene som var felles for PMTO og MST studiene presentert under avsnittet om PMTO. Dette var CBCL, YSR og TRF (Achenbach, 1991) og Family Satisfaction Scale (Lubrecht, 1992). Instrumentene som presenteres nedenfor ble bare benyttet i MST studien.

Plassering utenfor hjemmet: Ved inntak, ved avslutning av behandlingen og ved oppfølgingen to år etter rapporterte de foresatte i hver familie om: (1) hvor ungdommen bodde ungdommen på evalueringstidspunktet, og (2) hvor ungdommen bodde mesteparten av

de siste seks månedene. Svaralternativene var ”bodde hjemme” eller ”plassert utenfor hjemmet”. Ved oppfølgingen etter 2 år kunne de eldste ungdommene ha etablert seg på egen hånd (de eldste var da 19 år), og de som bodde for seg selv under tilsyn og/eller med samtykke av sine foreldre (eller Barnevernstjenesten) ble inkludert i gruppen ’bor hjemme’.

Ungdommenes egenrapportering. ”The Youth Self report” (YSR, Achenbach, 1991) har den samme oppbygningen som CBCL og TRF, med skåreverdiene 0-2 og delskalaer for internalisert og eksternalisert problematferd.

Selvrapportert kriminalitet og antisosialitet. ”The Self-Report Delinquency Scale” (SRD; Elliott, et al., 1983) ble fylt ut av ungdommene og målte hvor ofte de utførte antisosiale handlinger (f. eks., tyveri, salg av narkotiske stoffer, hærverk) ($\alpha = .95$).

Familieprosesser – gjensidig tilpasning og samhold. Ved begynnelse og avslutning av behandlingen fylte ungdom og foresatte ut ”Family Adaptability and Cohesion Evaluation Scales–III” (FACES-III, Olson, Portner, & Lavee, 1985) som måler graden av gjensidig tilpasning og samhold i familien.

Sosial kompetanse. Ungdommens sosiale ferdigheter i forhold til jevnaldrende ble målt med ”Social Competence with Peers Questionnaire” (SCPQ, Spence, 1995) og ble fylt ut av ungdommene, de foresatte og lærerne (α 's fra .81 til .93.). Utvalgte items fra ”The Social Skills Rating System” (SSRS, Gresham & Elliott, 1990) som også måler sosiale ferdigheter i forhold til foreldre og andre voksenpersoner (for eksempel lærere) ble brukt som supplement. SSRS leddene ble vurdert av ungdommene, de foresatte og lærerne og viste moderat reliabilitet ($\alpha = .64$).

Bruk av rusmidler. Ungdommenes alkohol og rusmisbruk ble målt med ”The Personal Experience Inventory (PEI: Winters & Henly, 1989). Ungdommene ble spurt hvor ofte de brukte alkohol og illegale rusmidler som cannabis. Skalaen består av 5 ledd som blir vurdert på en 4-punkts Likert skala fra 0 (aldri) til 3 (ofte). På begge tidspunkter var alpha koeffisienten 0.70.

Prosedyre

Ungdom og familier ble rekruttert til studien gjennom henvisninger fra den kommunale Barnevernstjenesten i de deltakende fylkene. Ungdommene og deres foresatte som var aktuelle for deltakelse i henhold til inklusjons- og eksklusjonskriteriene ble kontaktet for et hjemmebesøk. Familiene som på grunnlag av informasjon om studien ga skriftlig samtykke til deltakelse fylte deretter ut inntaksbatteriet. Kort tid etter hjemmebesøket ble familiene

informert om utfallet av randomiseringen og hvilket tilbud de ville få. På hvert utprøvningssted ble 25 saker tilfeldig fordelt på MST og ordinære tiltak med bruk av lukkede konvolutter som de fikk fra den forskningsansvarlige.

Inkludert i studien ble ungdom som 1) hadde begått lovbrudd, 2) viste annen antisosial atferd som aggressiv/voldelig atferd, hærverk, rømte hjemmefra og skulket skolen), 3) ruset seg i forbindelse med annen antisosial atferd, 4) hadde tilleggsproblemer som alvorlige vansker i skolen, dårlige sosiale relasjoner, verbal aggresjon, samt trusler om å skade andre, 5) var i alderen 13 til 17 år, og 6) hadde minst én foresatt som ungdommen bodde hos og som var tilstrekkelig involvert og motivert til å gjennomføre MST. Ungdom ble ekskludert fra evalueringsprosjektet hvis de: 1) ikke hadde en voksenperson i familien som kunne fungere som foresatt, 2) mottok annet tiltak, 3) hadde rusmisbruk som eneste problem, 4) var seksuelle overgrepere, 5) var diagnostisert med autisme, akutt psykose, eller var i fare for å begå selvmord, 6) bodde i et hjem der det var fare for ungdommens, familiemedlemmenes, eller behandlerens helse og sikkerhet, og 7) var involvert i en undersøkelsessak i den kommunale barnevernstjenesten.

Samtlige foreldre og ungdommer som ble forespurt, sa seg villige til å bidra, men fire familier trakk seg fra behandlingen før den ble iverksatt, og ble erstattet med fire nye familier. Når ingen ble ekskludert og få trakk seg før behandlingen startet kan det skyldes at inklusjonskriteriene var vide, men også at henvisende instans (de kommunale barnevernstjenestene) og familiene på forhånd hadde fått god informasjon om inklusjons- og eksklusjonskriteriene. De øvrige 55 ungdommene som ble rekruttert fra to behandlingssteder til studien av kjønnsforskjeller i MST gjennomgikk samme prosedyre som ungdommene i behandlingseffektstudien (for eksempel vurdert i forhold til inntaks og eksklusjonskriterier), men alle fikk Multisystemisk Terapi.

Behandlingsintegritet. Behandlingsintegritet ble målt gjennom oppringinger til familiene én gang i måneden og de besvarte 26 spørsmål om innholdet i møtene mellom familiene og MST terapeuten (TAMS: Treatment Adherence Measure Scores, Henggeler & Borduin, 1992, jf kapittel 5). Som tidligere nevnt viste det seg at et av utprøvningsstedene ikke etablerte prosedyren med å ringe opp familiene noe som gjorde det vanskelig å vurdere om det var MST behandling en evaluerte effekten av på dette utprøvningsstedet. Det finnes to måter å håndtere problemet med manglende iverksetting av TAM prosedyren på, enten å analysere forskjeller mellom utprøvningssteder med og uten TAM prosedyre eller å konsentrere analysene kun om utprøvningsstedene som har en tilfredsstillende rutine for måling av

programintegritet. Vi gjorde begge deler. I evalueringen etter behandlingsavslutning (Ogden & Halliday-Boykins, 2004) analyserte vi resultatforskjeller mellom de 4 utprøvningsstedene. Det viste seg at der man ikke hadde iverksatt måling av behandlingsintegritet, var MST resultatene dårligst. Ved de øvrige utprøvningsstedene viste det seg at høye skårer på behandlingsintegritet var forbundet med positive resultater. I oppfølgingsstudien 2 år etter inntak (Ogden & Amlund Hagen, 2006) begrenset vi analysene til de tre utprøvningsstedene som hadde gjennomført TAMS prosedyren. Det betyr at undersøkelsen ikke ble analysert i henhold til prinsippet om "intention-to-treat" (ITT), men som "treatment-of-the-treated" (TOT).

Intervensjon

MST-modellen er detaljert beskrevet i en publisert håndbok (Henggeler, m. fl., 1998; 2001; 2009). Behandlingsprogrammet ble implementert uten vesentlige modifikasjoner av originalen (jf kapittel 5). MST ble utført av 6 team og hvert team besto av 3-4 terapeuter og en veileder som også var teamleder. MST ble avsluttet når behandlingsmålene ble nådd i den enkelte sak og gjennomsnittlig behandlingstid var 25 uker (varierende mellom 7 til 38 uker). Henvisningene til prosjektet begynte kort tid etter at terapeutene hadde gjennomgått en ukes grunnopplæring og hadde avsluttet 3-5 pilot saker. MST Services i Charleston, South Carolina, en universitetstilknyttet opplæringsorganisasjon var ansvarlige for opplæringen av teamene og gjennomførte kvartalsvise vedlikeholdsseminarer for terapeutene. De ga også ukentlige telefon-konsultasjoner i alle saker (cirka 1 time pr. uke pr. team). Ungdommene i sammenlignings-gruppen og deres familier mottok de barnevernstiltak som de ville ha fått dersom MST ikke fantes. I denne undersøkelsen dreide det seg om langtidsplassering i institusjon, plassering i krise- og akuttinstitusjon for utredning samt oppfølging i hjemmet under veiledning av Barnevernstjenesten, eventuelt andre hjemmebaserte tiltak.

Statistiske analyser.

Pre-post MST studien hadde et 2 (MST vs US) x 2 (tid: før og etter behandlingen) x 4 (utprøvningssteder) mixed factorial design med randomisert fordeling på behandlingsbetingelser (Ogden & Halliday-Boykins, 2004). Kartleggingen ble gjennomført før familiene ble fordelt på behandlingsbetingelser og deretter ved avslutningen av behandlingen eller cirka 6 måneder etter inntak for sammenligningsgruppen.

For å beregne de mest reliable indikatorene på hver resultatvariabel og redusere Type I feil, ble det laget samlevariabler (eng. composite variables) basert på summen av

standardiserte variabler for internalisert problematferd (CBCL, YSR og TRF), eksternalisert problematferd (CBCL, YSR og TRF) og sosial kompetanse (vurderinger foretatt av foreldre, ungdommen og lærer der en summerte utvalgte ledd fra SCPQ, SSRS og CBCL). I tillegg beregnet vi middelværdien på råskårer for foreldrenes og ungdommenes vurdering av samhold og gjensidig tilpasning i familien (FACES).

I oppfølgingsstudien to år etter inntak ble resultatene analysert med kovarians analyser (ANCOVAs) og krysstabell (chi square) analyser (Ogden & Amlund-Hagen, 2006). Hvis inntaksskårene korrelerte med oppfølgingsskårene, ble de inkludert i analysene som kovariater for å kontrollere for eventuelle grupperforskjeller ved inntak og for eventuelle tidseffekter. Øvrige kovariater var alder og kjønn. Videre ble regresjonsanalyser benyttet for å undersøke hvor mye av variansen i resultatvariablene som ble forklart av behandlingsbetingelsene og kovariatene.

I studien av kjønnsforskjeller i MST ble gutter og jenters bosted sammenlignet med bruk av krysstabell analyser (chi square) og enveis variansanalyser (foreldre-, egen- og lærervurderinger) (Ogden & Amlund-Hagen, 2009). Fire multivariate analyser av kovarians (MANCOVA) ble gjennomført for å undersøke kjønnsforskjeller i atferdsendringer etter avsluttet behandling. MANCOVA reduserer sjansene for Type I feil som kan oppstå når en foretar mange signifikanstester på det samme utvalget. For å unngå Type I feil ble også signifikansnivået hevet til .01 på de øvrige analysene. Modell 1 undersøkte foreldrerapportering av internaliserende og eksternaliserende problematferd samt oppmerksomhetsproblemer (CBCL), modell 2 omfattet selvrapportering på de samme delskalaene (YSR), og i tillegg bruk av rusmidler og lov- og regelbrudd (delinquency) mens modell 3 undersøkte lærervurderinger på de tre førstnevnte delskalaene (TRF). Den fjerde modellen analyserte kjønnsforskjeller i foreldrenes evaluering av behandlingskvaliteten.

Missing data i oppfølgingsstudien. Fire familier trakk seg fra studien før den avsluttende kartleggingen (3 ordinære tiltak og 1 MST) slik at frafallet ved avslutningen av behandlingen var 4%. Gruppeforskjellene i sivilstatus mellom gruppene ved inntak viste seg ikke å ha noen betydning for behandlingsutfallet. Fra inntak til oppfølgingstidspunktet hadde 6 familier trukket seg fra prosjektet (3 MST og 3 ordinære tiltak) ved de tre utprøvningsstedene som ble inkludert i analysene, noe som utgjorde en frafallsprosent på åtte. Familiene skilte seg imidlertid ikke fra de øvrige når det gjaldt demografiske kjennetegn eller skårer på resultatvariablene (Ogden & Amlund-Hagen, 2006).

Resultater

Blant ungdommene som bodde hjemme ved inntak til studien, hadde ungdom i MST gruppen bodd hjemme de siste 6 månedene i signifikant større grad enn ungdom i sammenligningsgruppen ved behandlingsavslutning (91% mot 58%) (Ogden & Halliday-Boykins, 2004). MST ved behandlingsslutt var forbundet med redusert internalisert problematferd og en marginal reduksjon av eksternalisert problematferd samt økt sosial kompetanse hos ungdommene som fikk denne behandlingen sammenlignet med de som fikk vanlige tiltak. Samlevariabelen for ungdommenes problematferd viste at det var en signifikant reduksjon i internalisert problematferd ($p = .03$) og en marginalt signifikant reduksjon i eksternalisert problematferd ($p = .07$) i favør av de som hadde fått MST. Noen resultater ble bare evaluert ved behandlingsavslutning, så som familieprosesser, ungdommenes sosial kompetanse og familienes tilfredshet med tilbudet. Svar fra ungdommen, foreldrene, og lærer ble kombinert i én variabel, og ungdom som mottok MST skåret signifikant høyere på sosial kompetanse enn ungdom i sammenligningsgruppen. Det var også en trend i retning av at familiene i MST opplevde tiltaket de mottok som mer positivt enn det familiene i kontroll gruppen gjorde; 8 av 10 familier ville anbefale MST til andre eller bruke det om igjen hvis behovet skulle oppstå ($p = .07$). Begge gruppene rapporterte høyere grad av samhold i familien etter endt behandling, men på denne variabelen var det ingen behandlingseffekt i favør av MST.

I oppfølgingsstudien to år etter inntak ble atferdsvurderingene fra foresatte, lærere og ungdom analysert hver for seg (Ogden & Amlund-Hagen, 2006). MST-ungdom rapporterte om en signifikant reduksjon i kriminalitet og antisosial atferd (SRD, Elliott m.fl., 1983) i forhold til ungdom som mottok de vanlige barnevernstjenestene. Det var imidlertid ingen signifikant behandlingseffekt i favør av MST på Youth Self Report (YSR, Achenbach, 1991), men en signifikant nedgang over tid for begge gruppene i utagerende atferd og i total problematferd. Foreldrene i MST-gruppen rapporterte en signifikant nedgang i samlet problematferd på Child Behavior Checklist (CBCL, Achenbach, 1991) i forhold til foreldrene i kontrollgruppen. Videre rapporterte foreldrene i MST-gruppen en signifikant reduksjon i internalisert atferd hos sine barn sammenlignet med foreldrene til ungdom som mottok ordinære tiltak. Det var signifikante tidseffekter på alle CBCL delskalaer, det vil si at de foresatte i begge grupper vurderte at ungdommene hadde færre utagerende, internaliserende, og generelle problemer, men nedgangen var større hos MST-ungdommene.

Signifikante reduksjoner i utagerende atferd, internaliserende problemer, og samlet problematferd ble rapportert av lærerne til MST ungdommene sammenlignet med kontroll-

gruppens lærere (TRF, Achenbach, 1991). Det var også en generell nedgang i lærervurdert problematferd på alle de tre TRF delskalaene for alle ungdommene i undersøkelsen, men igjen var nedgangen signifikant større for de ungdommene som hadde gjennomgått MST.

Effekt størrelser ble benyttet for å analysere klinisk signifikans, og i oppfølgings-evalueringen to år etter inntak hadde selvrapportert antisosialitet (SRD) en effektstørrelse på 0.26, foreldrerapportert problematferd (CBCL) en effektstørrelse på 0.50 og lærervurdert problematferd en effektstørrelse på 0.68. En annen indikator på praktisk eller klinisk signifikans er antall ungdommer som skårer i 'normalområdet' for CBCL. Ved inntak skåret 9% av MST ungdommene i normalområdet, ved avslutningen av behandlingen var prosenttallet økt til 38% og ved oppfølgingen etter to år skåret 54% i normalområdet. De tilsvarende tallene for sammenligningsgruppen var 14%, 21.4% og 39%.

Frafallsprosenten i undersøkelsen var også lav. Mens henholdsvis 4% og 8% av deltakerne trakk seg fra forskningsprosjektet etter 6 og 24 måneder, så trakk ingen av familiene seg fra MST behandlingen tidligere enn forutsatt i behandlingsplanene.

I studien av kjønnsforskjeller ble jenter henvist signifikant oftere for skjulte (eng. covert) problemer som å skule skolen og rømme hjemmefra (63%) mens guttene ble signifikant oftere henvist for åpne (eng. overt) problemer som lovbrudd (37%) og vold i hjemmet (38%) (Ogden & Amlund-Hagen, 2009). Det var også en signifikant tendens til at gutter oftere enn jenter problemdebuterte før de hadde fylt 10 år. Blant de henviste ungdommene i denne studien hadde 51% av guttene startet sin problemkarriere før de var 10 år mot 18% av jentene. Det var få kjønnsforskjeller i inntaksvurderingene av problematferd; forskjeller ble påvist i lærervurderinger av eksternalisert problematferd og oppmerksomhetsproblemer der guttene fikk høyere skårer og jentene rapporterte om mer internalisert problematferd og om dårligere samhold i familien.

Analysene av komorbiditet eller tilleggsproblemer ble utført gjennom en dikotomisering av skårene på CBCL og TRF delskalaene. Hvis en ungdom skåret ett standardavvik over middelerdien for sitt kjønn på minst to delskalaer ble dette registrert som komorbiditet. Det var ingen signifikante kjønnsforskjeller i komorbiditet selv om det var en høy forekomst av tilleggsproblemer.

Analysen av kjønnsforskjeller i atferdsendring etter avsluttet behandling viste en positiv effekt av den multivariate analysen (MANCOVA) av foreldrevurderinger. Analyser av enkeltvariabler viste at jentene reduserte sin eksternaliserte problematferd mer enn gutter, mens det ikke var noen forskjeller i foreldrenes vurderinger av endring i internaliserte

problemer og oppmerksomhetsproblemer. Ungdommenes selvrappoteringer ga signifikant utslag på den multivariate testen og separate analyser viste at jentene rapporterte signifikant større reduksjon i lov- og regelbrytende atferd (delinkvens) mens guttene hadde signifikant mindre internalisert problematferd. Det var imidlertid ingen signifikante forskjeller i endring av eksternalisert problematferd eller oppmerksomhetsproblemer.

Den multivariate analysen av lærerrapporteringer ga ikke signifikant utslag, noe som tyder på at ifølge denne respondentgruppen endret gutter og jenter seg like mye i løpet av behandlingen. I den siste MANCOVA modellen ble ungdommenes og foreldrenes vurdering av samhold i familien testet, men den ga ikke signifikant forskjellig utslag for gutter og jenter. Videre viste analysene at plasseringer utenfor hjemmet var høyst uvanlig ved avslutningen av behandlingen for ungdom som hadde mottatt MST, og det var ingen kjønnsforskjeller. Det var heller ingen forskjell i foreldrenes vurdering av kvaliteten av behandlingstilbudt.

Studie nr. 5: Implementering og evaluering av evidensbaserte behandlinger av alvorlige atferdsproblemer hos barn og unge i Norge

Mens effektstudier tester om behandlingsprogrammer er virksomme i forskningspraksis og effektivitetsstudier om de virker i ordinær praksis, så retter implementeringsforskningen fokus mot spredningen av evidensbaserte programmer og metoder. Implementeringsforskning analyserer hvordan evidensbaserte intervensjoner kan formidles til praksisfeltet med en høy grad av program- og behandlingsintegritet. Denne siste studien i avhandlingen presenterer en begrepsmodell for implementeringen av evidensbaserte behandlingsprogram i arbeidet med alvorlige atferdsproblemer blant barn og unge i Norge. Den gir en oversikt over prinsipper og strategier som ble anvendt i implementeringen over en periode på åtte år, og gir samtidig et sammendrag av forskningen om MST og PMTO. Studien kan karakteriseres som en kasusstudie og en post-hoc analyse av implementeringsprosessen. Modellen oppsummerer erfaringene med implementering av evidensbaserte behandlingsprogram i Norge (Ogden, Amlund-Hagen, Askeland & Christensen, 2009) og består av følgende elementer: Programutvikling, formidling, adopsjon, implementering, tiltaksintegritet, lokal tilpasning og evalueringsresultater. Begrepsmodellen i figur 1 illustrerer forholdet mellom de ulike komponentene.

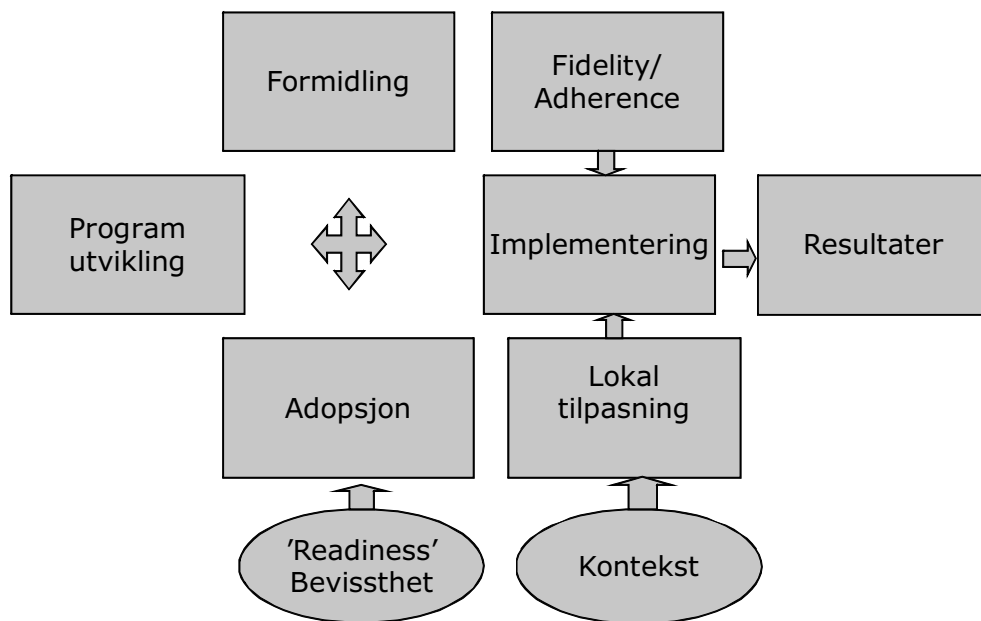


Fig 5.1 Implementeringsmodell (Ogden mfl., 2009).

Modellen har en pedagogisk funksjon gjennom å gi en oversikt over ulike implementeringskomponenter og forholdet mellom dem. Forholdet mellom komponentene er ikke linjært eller kronologisk, men illustrerer de gjensidige relasjonene mellom dem og hvordan de samlet eller komplementært påvirker resultatene.

Programutvikling

Siden det her dreide seg om behandlingsprogrammer som var utviklet i USA, er det lite å tilføye utover det som er skrevet om temaet i andre kapitler. De to miljøene som hadde utviklet programmene som ble tatt i bruk i Norge hadde i utgangspunktet svært forskjellig formidlingskapasitet. Som nevnt hadde MST både utarbeidet en behandlingsprotokoll eller håndbok (Henggeler mfl. 1998), en organisasjonsmanual (Strother, Swenson & Schoenwald, 1998) og etablert en velfungerende implementeringsorganisasjon (MST-services). Oregon miljøet som utviklet PMTO hadde ingen tilsvarende implementeringsorganisasjon i ryggen, og programutviklerne la ned et stort arbeid i å utvikle funksjonelle opplærings- og veiledningstiltak samt en behandlingsprotokoll i samarbeid med det norske implementeringsteamet for barn ved Atferdssenteret.

Formidling

Den nasjonale implementering av MST og PMTO dreide det seg om transport av behandlingsprogram utviklet i USA til Norge. I det norske implementeringsprosjektet kom den *linjære* modellen til å dominere, det vil si i hovedsak om en enveis formidling av kunnskap og kompetanse fra programutviklerne til praksismiljøer i Norge. Samtidig var det en omfattende kontakt mellom programutviklerne og de norske aktørene som omfattet Barne- og familiedepartementet og ansatte ved Atferdssenteret slik at en også kan snakke om en *relasjonsmodell* der kunnskap og erfaringer utveksles mellom programutviklere og de som implementerer (Ogden mfl., 2008; Ogden mfl., 2005).

Adopsjon

Adopsjon dvs beslutninger på organisasjonsnivå om å satse på et bestemt program ble i Norge tatt som en politisk beslutning etter tilråding fra to av departementene og om beslutninger tatt av fylkeshelsesjefene i 17 av 19 fylker om å implementere MST og PMTO. Forhold som har vist seg å fremme praktisk bruk av forskning er knyttet til stikkord som kvalitet, relevans og 'timing' (Nutley mfl., 2007) noe som passet bra i forhold til det norske prosjektet.

Forskningen om MST og PMTO var solid, og ble presentert på et tidspunkt hvor det var både politisk vilje og interesse i praksisfeltet for å prøve nye veier i arbeidet med alvorlige atferdsproblemer i barnevernstjenesten og psykisk helsevern. 'Timing' handler om hvorvidt forskningsresultatene presenteres på et gunstig tidspunkt, for eksempel når politiske beslutninger eller faglige veivalg skal tas om prioriteringer og videre satsing (Nutley mfl., 2007). Det forelå både grunnforskning om aggresjonsutvikling hos barn og unge (Patterson, 1982; Patterson mfl., 1992) og behandlingseffektstudier (Henggeler mfl., 1998; Patterson & Forgatch, 1995). Forskerne som sto for programutviklingen samt deres forskningsmiljøer hadde stor troverdighet selv om det ble reist noen innvendinger mot de kommersielle aspektene ved MST. Representanter for helsedirektørene i alle landets 19 fylker invitert til Oslo og spurt om de ville medvirke i implementeringen av de nye programmene.

Departementene finansierte opplæring, veiledning, praktisk støtte og forskningen mens de lokale tjenestene lønnet praktikerne og betale deres utgifter i forbindelse med opplæring og praktisering av programmene. I noen grad og med betydelige variasjoner i 'readiness' takket representantene ja til tilbudet, noe som har ført til variasjoner i tettheten av PMTO terapeuter og MST team i Norge. Men siden praktisk talt alle fylkene valgte å delta i utprøvingen av programmene, gikk en umiddelbart i gang med en nasjonal implementering i relativt stor skala.

Blant Rogers (1995) kriterier for vellykket adopsjon var det sannsynliggjort at de nye intervensjonene hadde *relative fordeler* i forhold til eksisterende praksis. Det var stor etterspørsel i praksismiljøene etter nye metoder og tjenester i arbeidet med alvorlige atferdsvansker, og de forskningsresultatene fra USA virket troverdige. Mulighetene var også tilstede for at en kunne *observere* de nye programmene i bruk, og norske fagfolk reiste på studiebesøk både til USA og Canada for å sette seg nærmere inn i programmene (Eugene i Oregon, Charleston i South Carolina og London i Canada). Vanskeligere var det på forhånd å vurdere om intervensjonene var *kompatible* eller forenlige med eksisterende praksis og deres *kompleksitet*.

Implementering

Den nasjonale implementeringsstrategien kan beskrives som en policy basert på langsiktig finansiering og formidling av evidensbaserte tiltak til ordinær praksis og etableringen av et nasjonalt senter for implementering og forskning. Planene var allerede fra starten av å lokalisere MST til den fylkeskommunale barnevernstjenesten mens PMTO skulle praktiseres både i psykisk helsevern og barnevernstjenesten. Rutiner for forhåndsevaluering (eng. site assessments) av praksissteder ble introdusert av MST-services (Strother mfl., 1998) og terapeuter ble nyrekruttert fra flere yrkesgrupper (MST) eller rekruttert internt fra det ordinære tjenestesystemet. På denne måten ble de evidensbaserte programmene implementert på tvers av tjenestesystemene for barn og ungdom og det ble etablert samarbeidsforbindelser innenfor programmene, men på tvers av profesjoner og tjenester. Det ble registrert en betydelig interesse blant kvalifiserte praktikere for de utlyste stillingene. Videre ble det lagt ned en betydelig innsats i arbeidet med å formidle kunnskap om de nye behandlingsprogrammene gjennom fagtidsskrifter, konferanser og møter. Blant annet ble Atferdssenterets årlige fagkonferanse et viktig forum for spredning av kunnskap gjennom introduksjonsseminar og rapporter om forskningsresultater. I tillegg viste media interesse for det nye tilbudet og dets muligheter.

Implementeringsstrategien kan beskrive som en kombinasjon av 'bottom-up' og 'top-down' strategier (Ferrer-Wreder mfl., 2004). 'Bottom-up' eller 'nedenfra-og-opp' strategien beskriver en desentralisert tilnærmingen til utvikling av ny praksis, der lokale tjenester eller den enkelte praktikere ut fra egne behov velger arbeidsmodeller, metoder eller program. Dette er den dominerende utviklingsstrategien i barne- og ungdomstjenestene og støtter opp under prinsippet om faglig autonomi og fritt metodevalg. Fordelen med denne strategien er at praktikerne utvikler et lokalt eierforhold til arbeidsmåten, noe som øker motivasjon og grad

av forpliktelse. Ulempen kan være at det er vanskelig å opprettholde tilbudet med tilstrekkelig omfang, intensitet og kvalitet over tid. Utskifting av personale og vekslende prioriteringer blant de ansatte kan føre til at man stadig skifter tilnærming eller at kvaliteten i tilbudet synker over tid. Når det var interesse og støtte i praksisfeltet for implementeringen av evidensbaserte behandlingsprogram i Norge, henger det sammen med at ansatte i barne- og ungdomstjenestene hadde gjort seg kjent med slike tilnærminger gjennom litteratur, konferanser og kontakt med internasjonale fagmiljøer (Fossestøl & Enehaug, 2008). Dette kom blant annet til uttrykk gjennom den store pågangen og interessen for å delta i opplæringsprogrammet for PMTO og antall søkere til stillinger som MST terapeuter. Positive resultater og tilbakemeldinger fra de første familiene som mottok behandlingstilbudene bidro også til å øke interessen blant praktikerne.

En 'top-down' eller 'ovenfra-ned' modell er en sentralisert utviklingsstrategi der beslutninger fattet på sentralt nivå i politikk eller forvaltning blir formidlet til lokalt nivå for å omsettes til praksis. Fordelen med denne strategien er at en kan samordne ressurser og innsats og dermed tilby lokale tjenester opplæring, veiledning og praktisk støtte. Ulempen består i at strategien kan mobilisere motkrefter lokalt, hvis prioriteringene ikke stemmer overens med lokale behov eller hvis praktikerne føler at deres profesjonelle autonomi trues. Tidligere erfaringer med sentrale initiativ hadde vist at disse kan ha begrenset gjennomslagskraft hvis de virker lite relevante og attraktive for praktikere og lokale tjenester. Det var betydelig sentral støtte til implementeringen av PMTO og MST, først og fremst fra Barne- og Familiedepartementet, men også fra andre departementer som ga økonomisk støtte til implementeringen. Det var også en tydelig interesse for å etablere programtilbudene i den fylkeskommunale barnevernstjenesten og senere i det statlige barnevernet (Barne- og Familiedirektoratet og BUF-etat). En viktig del av 'top-down' strategien var Atferdssenterets etableringen av prosedyrer og rutiner for opplæring, veiledning og kvalitetssikring av praksis. I PMTO ble det etablert et 18 måneders omfattende opplæringsprogram som førte fram til sertifisering som PMTO terapeut. Sertifiseringen var basert på video-opptak av terapeutenes behandlingstimer som ble vurdert av PMTO eksperter ved Atferdssenteret og fornyes hvert tredje år. Andre tiltak som skal bidra til å sikre høy grad av program- og behandlingsintegritet i PMTO er kontinuerlig utvikling av håndbøker og praktisk materiell, og løpende veiledning i regionale grupper. På tilsvarende måte er det etablert et omfattende system for kvalitetssikring av MST med løpende registrering av TAM skårer og ekspertkonsultasjon i alle saker, monitorering av stabilitet i behandlingsresultater etter 6, 12 og 18 måneder, samt regelmessige lokale evalueringer av produktivitet og behandlingstid. Hensikten med disse omfattende

prosedyrene var å sikre høy grad av program- og behandlingsintegritet og å forebygge avdrift fra programmet.

Implementeringen av MST og PMTO ble organisert som en kombinasjon av ”top-down” og ”bottom-up” strategiene slik at programmene ble adoptert basert på råd fra en ekspertgruppe, men implementert på basis av støtte fra den nasjonale og fylkeskommunale forvaltningen. På norsk side ble oppgavene knyttet til utvikling, implementering og forskning først ivaretatt av Atferdsprosjektet ved Universitetet i Oslo og fra 2003 av Atferdssenteret (Norsk senter for studier av problematferd og innovativ praksis). Hensikten med sentret var å knytte forskning, implementering og praksis tettere sammen. Forhold som virket fremmende på implementeringen av PMTO og MST var en generell bekymring for underkapasitet og manglende kvalitet på behandlingstilbudet til barn med alvorlige atferdsproblemer, en generell etterspørsel etter og interesse for evidensbaserte behandlingsmetoder blant praktikere, og en vilje i forvaltningen til å finansiere og støtte implementeringen av behandlingsmetodene. Positive tilbakemeldinger fra familier som hadde gjennomgått behandlingen var også som nevnt medvirkende til at mange familier ble henvist eller tok kontakt for å få disse behandlingstilbudene. Det kom også positive signaler om å styrke familiebaserte tiltak fra det politiske miljøet, og en uttrykt vilje til å satse på tidlig identifikasjon og intervensjon i forhold til atferdsvansker.

7. Avsluttende drøfting

Den norske PMTO studien viste at foreldrebehandling kan ha en signifikant positiv effekt på barns problematferd målt ved avslutningen av behandlingen i et klinisk utvalg. Positive hovedeffekter ble målt for foreldrevurderinger av barnas samlede problematferd og eksternaliserte problemer (CBCL) men også for foreldrenes daglige rapporteringer over tre dager (PDR). PMTO foreldrene ga også mer positive vurderinger av behandlingstilbudet. I følge lærervurderingene forbedret barna i PMTO gruppen sine sosiale ferdigheter signifikant i løpet av behandlingen sammenlignet med barna i sammenligningsgruppen, noe som tyder på generaliseringseffekter fra hjem til skole. I diskusjonen om evidensbasert praksis har problemstillingen om programmer virker eller ikke virker blitt supplert med spørsmålet om hva som virker for hvem og hvor bra. Det handler om hvilke forhold eller variabler som modererer behandlingsresultatene, som for eksempel alder, kjønn og problembelastning. Blant interaksjonseffektene i denne studien var alder den tydeligste moderatoren og viste at resultatene var signifikant mer positive for PMTO enn for kontrollbetingelsene på flere variabler i både foreldre- og lærervurderinger for barn under åtte år. Det ble også funnet interaksjonseffekter mellom kjønn og behandling, men siden antall jenter var begrenset i denne som i tilsvarende studier av foreldreprogrammer (Lundahl mfl., 2006), må estimatet betraktes som mindre robust. Det er imidlertid gjort undersøkelser med større utvalg som støtter antakelsen om at jenter kan ha et like stort utbytte av PMTO som gutter (Kjølbi & Ogden, 2009).

Resultatene ga begrenset støtte til PMTO modellen for foreldreferdigheter slik at bare én av fem ferdighetsgrupper viste signifikante endringer i PMTO foreldrenes favør. Det var en signifikant økning i PMTO foreldrenes ferdigheter i *grensesetting*, og dette er gruppen av foreldreferdigheter som ved siden av tilsyn har vist seg å være den mest sentrale i arbeidet for å endre antisosial atferd hos barn (Patterson mfl., 1992; Patterson, 2002). Resultatene fra analysen av stimodellen indikerte at foreldrenes grensesetting var en endringsmekanisme for barns atferd, noe som understøtter tidligere forskning om denne sammenhengen (Forgatch, 1991). Men som nevnt i kapittel 4 burde foreldrenes ferdigheter i grensesetting (mediatoren) ha blitt målt på et tidligere tidspunkt enn atferdsendringene hos barna for å påvise en medierende effekt (Kazdin & Nock, 2003; Weersing & Weisz, 2002). I stianalysen ble

endringsmekanismen (grensesetting) og resultatvariablene (barnas atferd) målt på samme tidspunkt, noe som gjør det vanskelig å avgjøre i hvilken retning påvirkningene gikk.

Resultatene fra den norske PMTO studien understøttes av tidligere forskning som har vist at atferdsorienterte foreldreprogram gir signifikant positive atferdsendringer, men med moderate effektstørrelser. Videre underbygger resultatene konklusjonene som går ut på at atferdsorienterte foreldreprogram gir bedre resultater for yngre enn for eldre barn og at høy behandlingsintegritet har positive effekter. Når studien ikke bekreftet hele modellen for foreldreferdigheter kan det henge sammen med at ikke alle ferdighetsgruppene ble formidlet på en virksom måte i behandlingene, men det er også mulig at målingene ikke var optimale. Foreldreferdighetene som ble kodet på grunnlag av kodernes globale inntrykk fra strukturerte samspill situasjoner kan ha vært utilstrekkelige for å kartlegge foreldrenes ferdigheter og burde ha blitt supplert med hjemmeobservasjoner over et lengre tidsrom eller med selvrapportering fra foreldrene.

Blant PMTO studiens sterke sider bør det nevnes at behandlingsprogrammet ble testet i ordinær praksis og satt på en stringent prøve gjennom sammenligningen med andre aktive behandlingstilbud i en randomisert kontrollert studie. Det er også grunn til å bemerke at det var lite frafall i behandlingsperioden og at det ble påvist positive generaliseringseffekter til sosiale ferdigheter i skolen. Begrensningene knytter seg særlig til at beskrivelsen av behandlingstilbudet til sammenligningsgruppen var ufullstendig. Det er også grunn til å påpeke at selv om samhandlingen i familien ble analysert gjennom grundige mikrososiale koder før og etter behandlingen, så bidro disse dataene i liten grad til å kaste lys over behandlingsresultatene.

Evalueringen av MST i Norge viste at programmet var mer virksomt enn tjenestetilbudet som vanligvis ble gitt til ungdommer med alvorlige atferdsproblemer som ble henvist til Barnevernstjenesten (Ogden & Halliday-Boykins, 2004). Både i evalueringen umiddelbart etter avslutning av behandlingen og i oppfølgingsstudien 2 år etter inntak, viste resultatene at MST ungdommene hadde hatt en mer positiv utvikling enn de som hadde fått andre tilbud og annen oppfølging fra Barnevernstjenesten (Ogden & Amlund-Hagen, 2006). Det er særlig grunn til å framheve at MST i høyere grad enn andre tiltak hadde forebygget at ungdom ble plassert utenfor hjemmet. I oppfølgingsstudien rapporterte foreldre og lærere i MST gruppen også om signifikant mindre eksternalisert problematferd og ungdommene om færre lov- og regelbrudd enn informantene i kontrollgruppen. Når resultatene var inkonsistente for ungdommenes selvrapportering (YSR vs. SRD) kan det tyde på at MST hadde en målrettet

effekt på problematferden som behandlingen rettet seg mot, mens effekten var mindre på problematferd i vid forstand. Samtidig viste det seg noe uventet at foreldrene og lærerne i MST gruppen rapporterte mindre internalisert problematferd enn foreldre og lærere i sammenligningsgruppen.

Selv om forskjeller mellom behandlingssteder (eng. site differences) ikke var noe sentralt tema da undersøkelsen ble planlagt, så viste post-hoc analyser at det var forskjeller i behandlingsutbytte mellom ungdommene på de fire utprøvningsstedene. Disse forskjellene tydet på at manglende implementering av prosedyrer for måling av behandlingsintegritet ga de dårligste resultatene mens for de øvrige utprøvningsstedene var det en positiv korrelasjon mellom terapeutenes behandlingsintegritet (TAM skårer) og ungdommenes behandlingsutbytte.

Resultatene fra den norske MST evalueringen var sammenfallende med forskningsresultatene fra studier i USA. MST framsto som en robust og fleksibel behandlingsform som kunne tilpasses ungdom med varierende kjennetegn, og alder og kjønn var svake moderatorer. Og på samme måte som i PMTO studien ble det påvist generaliseringseffekter til skole og til foreldre- og lærervurderinger av ungdommenes internaliserte problematferd.

Undersøkelsen aktualiserte også spørsmålet om betydningen av forskjeller mellom intervensjons- og kontrollbetingelser i Norge sammenlignet med USA. Tilbudet til kontrollgruppen i Norge var mer omfattende og behandlingsorienterte enn tilbudet til amerikanske kontrollgrupper. Tidligere MST studier i USA rapporterte om sammenligningsgrupper som besto av ungdomskriminelle (eng. juvenile delinquents) som var under tilsyn (eng. probation office visits) og som bare ble henvist til sosialtjenesten hvis det ble vurdert som helt nødvendig. I den norske studien fikk sammenligningsgruppen aktive tilbud som omfattet institusjonsplassering og hjelpetiltak i regi av barnevernet, inkludert hjemmebaserte tiltak. Dette kan forklare at den relative effekten av MST i Norge var lavere enn i USA, selv om resultatene i Norge viste at MST ga bedre resultater enn relativt godt utviklede barnevernstiltak. De norske effektstørrelsene var mindre enn de som ble funnet i effektstudier i USA, men på høyde med amerikanske effektstørrelsene når MST ble testet i ordinær praksis.

Blant MST studiens sterke sider bør det framheves at de positive resultatene ble oppnådd i en randomisert kontrollert effektivitetsstudie (dvs. gjennomført i ordinær praksis) der MST ble sammenlignet med forsvarlige og ofte omfattende andre barnevernstiltak. Det er også grunn til å framheve at forskjellene i MST gruppens favør økte fra avslutningen av behandlingen til oppfølgingstidspunktet. Deltakelsesgraden var også høy hvis en ser bort fra utprøvningsstedet som ble utelatt i oppfølgingsstudien. Både i denne og PMTO studien bør det

også nevnes at evalueringene ble gjennomført av et uavhengig forskerteam. Det innebærer at forskerne ikke hadde kontakt med familiene i løpet av evalueringen, de var heller ikke ansvarlige for opplæring eller veiledning av terapeutene. I den grad forskerne kan sies å ha hatt preferanser, så knyttet disse seg til evidensbaserte intervensjoner og ikke til enkeltprogrammer. Dette tilsvarer det som har blitt omtalt som ”multi-allegiance” forskerteam.

Begrensningene ved studien var for det første at ett av utprøvningsstedene ikke iverksatte rapporteringssystemet for behandlingsintegritet (TAM). Når dette stedet ble ekskludert fra analysene bryter det med prinsippet om ’intention to treat’. Samtidig førte ekskluderingen til en reduksjon i statistisk styrke slik at signifikanskravene til gruppeforskjeller økte. En annen begrensning var at i denne som i PMTO studien ble behandlingstilbudet til sammenligningsgruppen ufullstendig beskrevet med hensyn til innhold og foreløp. En tredje begrensning ved studien var den lave svarprosenten fra lærere ved oppfølgingstidspunktet, noe som kan forklares med ungdommenes ustabile skolekarrierer og at mange avsluttet skolegangen etter grunnskolen. Og selv om lærervurderingene stemte godt overens med undersøkelsens øvrige resultater og hypoteser, så blir det viktig i senere studier å undersøke ungdommenes prestasjoner og skolekarrierer mer inngående.

Interaksjonsanalysene tydet på at alder og kjønn hadde en modererende effekt på noen få variabler; MST motvirket plasseringer utenfor hjemmet bedre for gutter enn for jenter og bedre for eldre (> 17 år) enn for yngre ungdommer. Og siden det var få jenter ($N = 27$) med i analysene var det behov for å se nærmere på alders- og kjønnsforskjeller i behandlingsutbytte i et utvalg som hadde en høyere jenteandel (se Ogden & Amlund-Hagen, 2009). En utvidet analyse viste at alder ikke var noen viktig moderator for ungdommene i denne aldersgruppen (Otnes & Prydz, 2009). Undersøkelsen om kjønnsforskjeller i MST viste at selv om jentene hadde en noe annen problemprofil enn guttene ved inntak, så endret de seg like mye som guttene i løpet av behandlingen (Ogden & Amlund-Hagen, 2009). Jentene som ble henvist til MST var sammenlignet med guttene like problembelastet med hensyn til utagerende atferd og rusmisbruk, de hadde like mange tilleggsproblemer, men de startet oftere sin problemkarriere på et senere tidspunkt. Selv om gjennomsnittsalderen for henvisning til MST var den samme for begge kjønn (15 år), var det klare indikasjoner på at flere gutter enn jenter hadde utviklet problemer før 10 års alderen. Resultatene støttet derfor ikke ”gender-paradox” hypotesen som går ut på at når jenter først henvises eller diagnostiseres med ’atferdsforstyrrelse’, så er de mer problembelastet enn guttene (Eme, 1992). Styrken ved denne studien var at den inkluderte en

større andel jenter enn behandlingseffektstudien, og at det derfor var mulig å gi en mer utfyllende beskrivelse av kjønnsforskjeller i inntaks-kjennetegn og av endring under behandlingen. Men med det foreliggende undersøkelses-designet var det ikke mulig å sammenligne resultatene på tvers av kjønn og behandlings-betingelser med en kontrollgruppe som fikk andre tilbud. Det eneste som derfor kan fastslås var at jentene endret seg like mye som guttene, og at MST var en virksom behandling for de fleste ungdommene i utvalget. Studien kunne heller ikke besvare spørsmålet om det var forskjeller i behandlingsinnholdet for jenter og gutter, og om likhetene i behandlingsutbytte varte over tid.

Samlet viste studiene at MST og PMTO som empirisk støttet behandlingsprogram ga bedre resultater enn ordinære tjenester for målgruppene. Viktige fellestrekk ved programmene var at de hadde et teoretisk fundament, et strukturert opplærings og veiledningstilbud, en definert behandlingsprotokoll samt målinger av behandlingsintegritet og behandlingsutfall. Begge undersøkelsene hadde flere indikasjoner på positive generaliseringseffekter fra familie-behandling til barnas skoleatferd observert av lærere.

Studiene av MST og PMTO representerte effektivitetsstudier, dvs evaluering av program-effekter når de ble implementert av kompetente terapeuter i ordinær praksis. Men fordi programmene ble testet i ordinær praksis hadde forskerne mindre kontroll med rekruttering av terapeuter og klienter og med selve gjennomføringen av behandlingen. Det kan ha ført til at utvalget av terapeuter og klienter var mindre homogent enn det som finnes i effektstudier og at behandlingene kan ha bli påvirket av kontekstuelle forskjeller mellom behandlingsstedene. Fordelene var imidlertid knyttet til generalisering av resultatene, da utvalget av terapeuter og klienter antagelig var ganske representativt for de som tilbyr og mottar slik behandling i Norge.

Betydningen av implementering med høy programintegritet, behandlingsintegritet og kompetanse går som en rød tråd gjennom forskningen om evidensbaserte behandlings-programmer. Tiltak bør iverksettes og gjennomføres som planlagt eller i samsvar med programmenes retningslinjer. Det handler om å få tiltak til å virke og redusere avstanden mellom idealer og virkelighet, eller mellom teori og praksis. Målene på tiltaksintegritet i MST (Treatment Adherence Measure Scores, TAMS) og PMTO (Fidelity of Implementation System, FIMP) ble gjennomført på foreskrevet måte og viste at det var en sammenheng mellom gjennomsnittsskårene for behandlingsstedet og behandlingsutfall for MST og sammenheng med foreldreferdigheter og tilfredshet med behandlingen i PMTO. Selv om resultatene pekte i forventet retning, var mangelfull gjennomføring av prosedyrer et problem,

slik at det ble få slike målinger. Det førte som nevnt til at ett utprøvningssted falt ut i MST studien, og at antall video opptak som lå til grunn for analyser av integritetsmålinger i PMTO var lavt. I senere forskningsprosjekter bør det derfor være en prioritert oppgave å sikre at integritetsmålinger blir gjennomført systematisk og konsekvent i alle behandlinger og for alle terapeuter.

De norske evalueringene av PMTO og MST hadde primært til hensikt å demonstrere hovedeffekter av programmene, at de var relativt mer resultateffektive enn ordinær praksis. Forholdet mellom programfaktorer og terapeutiske fellesfaktorer var derfor ikke noe sentralt tema i undersøkelsene, selv om det har vært framtrедende i debatten om virksomme faktorer i psykoterapi. Det finnes utvilsomt terapeutvariasjoner innenfor rammene av de aktuelle programmene, og det er også sannsynlig at andre terapeutiske fellesfaktorer har påvirket resultatene. Men fellesfaktorene har i disse undersøkelsene blitt sett på som en integrert del av behandlingstilbudet og at forholdet mellom terapeuter og program har vært et spørsmål om 'både-og', heller enn 'enten-eller'. I noen grad ble fellesfaktorene testet i interaksjons- eller moderator analyser, men langt fra på noen utfyllende måte. I senere analyser kan det derfor være aktuelt å se nærmere på betydningen av terapeutiske fellesfaktorer for behandlingsutbytte. Men på bakgrunn av det som her er nevnt, bør det ikke føre til en prinsippdebatt hvor programutvikling settes opp mot utvikling av terapeutenes kompetanse. Erfaringer og resultater fra de norske evalueringstudiene tyder på at det er mulig å anvende klinisk skjønn og kreativitet innenfor rammene av evidensbaserte behandlingsprogrammer, og at verken bruk av behandlingsprotokoller eller måling av behandlingsintegritet hadde noen negativ effekt på terapeutenes kompetanse eller på behandlingsresultatene. De norske studiene viste også at kritikken om at evidensbaserte intervensjoner bare dreide seg om korttidsbehandling av homogene målgrupper ikke traff helt. Både i PMTO og MST var avslutningen av terapiene kriteriebasert heller enn tidsbestemt, og det var mulig å forlenge behandlingen hvis deltakerne ser ut til å ha nytte av det. Videre hadde begge programmene vide inklusjonskriterier, og blant tilleggsproblerne i de norske utvalgene inngikk ADHD, internaliserte problemer (angst og depresjon), skolevansker og for de eldste, rusproblemer.

Studiene syntes å bekrefte at MST og PMTO som ble utviklet i USA hadde transatlantisk relevans. I sprednings- og implementeringsprosessen krysset programmene både geografiske grensser og språkgrenser. Studiene har i et komparativt perspektiv illustrert interessante kontekstuelle forskjeller mellom Norge og USA, selv om ikke alle forskjellene er like godt dokumenterte. Men det er mye som tyder på at det er flere og mer alvorlige

miljømessige risikofaktorer i USA enn i Norge, så som fattigdom og tilgang på våpen og narkotiske stoffer. Samtidig er det betydelige likheter i de mikrososiale prosessene i familier, skoler og barnehager og som kan føre til opptrappende konflikter og relasjonsproblemer i familier samt bidra til at barn utvikler alvorlige atferdsproblemer.

Som tidligere nevnt var det den relative effekten av programmene som ble testet i de norske replikasjonsstudiene. Når effektstørrelsene gjennomgående var mindre i de norske enn i de amerikanske studiene kan det i noen grad forklares med forskjeller i måten en behandlet sammenligningsgruppene på. Det ser ut til at MST og PMTO førte til tilnærmet like positive endringer i norske familier som i familier i USA (se også Sundell mfl., 2008). Samtidig er det mulig at kontrollgruppene hadde en mer positiv utvikling i de norske enn i de amerikanske studiene blant annet fordi Norge har et mer utbygd og tilgjengelig offentlig tjenestesystem av rimelig god kvalitet. I sammenligningen av MST resultatene mellom Sverige og Norge pekte Löfholm mfl. (2009) på at implementeringen i Norge ble støttet økonomisk av Barne- og familiedepartementet og faglig av Atferdssenteret. Dette kan ha hatt en positiv virkning på behandlingsintegriteten og implementeringskvaliteten i Norge sammenlignet med Sverige der implementeringen ble gjennomført uten et nasjonalt støtteapparat. Löfholm mfl. (2009) mente også at man i Sverige manglet kunnskap om hvor virksomt det eksisterende tiltaksapparatet for ungdom med alvorlige atferdsproblemer var, før en implementerte MST. En slik kunnskap mente forfatterne var en viktig forutsetning for å kunne vurdere hvilke forbedringer en kunne forvente å få gjennom å implementere nye tiltak som MST. En grundigere forhåndsevaluering av ordinære tjenester kan derfor være nyttig før en sammenligner dem med nye intervensjoner. Det kan hende at ambisjonsnivået noen ganger skal senkes slik at målsettingen er å introdusere nye tiltak og metoder som gir minst like gode resultater som de som allerede finnes. Men det kan være et poeng at de skal være lettere tilgjengelige, settes inn på et tidligere tidspunkt i barnas liv, gjennomføres i barnas og de unges hjemmemiljø og dermed være tids- og kostnadsbesparende. Dette er aktuelle problemstillinger for framtidige forskningsprosjekter på området.

Det er behov for flere randomiserte kontrollerte studier av modell programmer som MST og PMTO. Det er viktig å undersøke om programmene opprettholder sin effektivitet over tid og i nye populasjoner, og resultatene av de gjennomførte studiene har ikke ubegrenset gyldighet. Det er også behov for å evaluere andre programmer og modifiserte varianter av PMTO og MST, for eksempel programmer av kortere og lengre varighet, eller å teste hvilke resultater en

får ved å legge til nye komponenter. Det er også behov for kontrollerte evalueringsstudier med større grupper, slik at en kan utføre mer robuste analyser av moderatorer og potensielle endringsmekanismer. Diskusjonen om den relative betydningen av spesifikke programfaktorer og terapeutiske fellesfaktorer bør følges opp med analyser der en tester den relative betydningen av terapeut- og klientkjenne tegn, behandlingsallianse samt ekstraterapeutiske forhold som sosial støtte og livshendelser. Analyser av sammenhengene mellom enkeltkomponenter og behandlingsutfall er også aktuelle. Det er behov for å undersøke om noen behandlingskomponenter eller ingredienser er særlig virksomme i arbeidet med å endre antisosial atferd.

Videre mangler det langtidsstudier av barn og unge som har gjennomgått PMTO og MST for å undersøke om de positive effektene vedlikeholdes i ungdoms- og voksen alder. Det kan særlig være aktuelt med registerstudier der en analyserer hvordan de klarer seg i utdanning og arbeidsliv, om og om de avholder seg fra å begå lovbrudd. Selv om det er gjennomført oppfølgingsstudier, så er disse av relativt kort varighet mens den underliggende hypotesen i disse programmene er at behandlingen skal ha en langsiktig effekt på atferd og samfunnsmessig tilpasning. Spørsmålet er hvordan det går med disse barna og ungdommene på lang sikt. Klarer de å styre unna kriminalitet og rusavhengighet, og klarer de å skaffe seg utdanning, holde på en jobb, stifte familie og bli omsorgsfulle foreldre?

Et viktig tema i intervensjonsstudier er endringsmekanismer, det vil si hvilke prosesser som medierer endring (LaGreca & Silverman, 2009). I PMTO har dette særlig handlet om foreldreferdigheter, og i den norske studien viste at behandlingen påvirket barnas problematferd gjennom foreldrenes grensesetting. I MST har man påvist at graden av kontakt med antisosiale venner og samhold i familien har fungert som endringsmekanismer, men den norske studien replikerte ikke slike analyser. Det er også behov for studier der en analyserer om andre endringsmekanismer enn foreldreferdigheter og familieprosesser er virksomme i behandlingstiltak, for eksempel prosesser knyttet til venner, partnere, skole/utdanning og lokalmiljø. Mye tyder også på at nye undersøkelser bør ha mer fokus på terapeutiske fellesfaktorer og at en i tillegg til programeffekter bør studere den modererende effekten av terapeutkjenne tegn, klientvariabler og forhold som livshendelser og sosial støtte.

Til slutt bør behovet for implementeringsstudier understrekes, der hensikten er å studere prosessen hvor evidensbaserte programmer formidles til aktuelle praksismiljøer og hvor godt disse lykkes med å implementere programmene på langsiktig basis. Slike studier kan ta sikte på å kartlegge i hvilket omfang man når fram til de som kan nyttiggjøre seg programmene og hvordan programmene innpasses i forhold til eksisterende praksis.

Behandlingsprogrammene som er tema for denne avhandlingen har blitt opprettholdt etter at effektivitetsstudiene var avsluttet og både PMTO og MST representerer i dag et utbredt behandlingstilbud til barn, unge og familier over store deler av landet. Ti år etter etableringen har MST og PMTO ved siden av "De utrolige årene" blitt dominerende program i arbeidet med å behandle atferdsproblemer blant barn og unge. Med de store investeringene som er gjort i opplæring, veiledning, forskning og kvalitetssikring er det mye som tyder på at programmene vil bli videreført i overskuelig framtid. Men etableringen har imidlertid ikke vært til hinder for at andre evidensbaserte program har blitt implementert og evaluert av andre aktører, så som De utrolige årene (Larsson mfl., 2009). Også ved Atferdssenteret er andre program etablert så som det skoleomfattende intervensjonsprogrammet PALS (Arnesen, Ogden & Sørli, 2006; Sørli & Ogden, 2007), og det forebyggende programmet "Tidlig Innsats for Barn i Risiko (TIBIR). Atferdssenteret er også involvert i utprøving av Funksjonell familierapi (FFT) og Multidimensional Treatment Foster Care (MTFC). Nye program kan derfor være komplementære eller alternativer til de som allerede er etablerte.

Selv om MST og PMTO har vist seg å være effektive intervensjoner, så var det familier som ikke responderte på behandlingen, og barn og unge som bare endret seg marginalt. Det er derfor en utfordring for framtidige studier å evaluere hvorfor noen barn, unge og deres familier hadde et begrenset utbytte, og hva som kan gjøres for å forbedre programmene. Blant utfordringene og utviklingsmulighetene inngår behovet for bedre å kunne styre klientstrømmen i kommunale tjenester og spesialisttjenestene slik at en bedre utnytter den kapasiteten som de nye programmene innebærer. Tiltakene bør også ha en større bredde slik at skoletiltak og fritidstiltak inkluderes. Ikke minst bør en vurdere hvordan en kan integrere familiebehandling med skoleintervensjoner, for eksempel gjennom undervisning i sosiale ferdigheter og støtte til sosial kompetanseutvikling. På foreldresiden ser det ut til at det bør arbeides mer med utvikling av pro-aktive foreldreferdigheter som gjør det mulig å endre barns atferd i konfliktfrie situasjoner og perioder. Det ser også ut til at barna i større grad bør involveres i behandlingen og delta i timene.

Norge har en lang tradisjon for å bruke internering og plasseringstiltak overfor barn og unge med alvorlige atferdsproblemer og gjennom årene har de blitt plassert i spesialskoler, verneskoler og ulike former for institusjoner (Befring utvalget, 2004; Kjellsberg & Dahl, 1998; NOU 1981). En har i denne sammenhengen snakket om eksporttendenser der barna har blitt flyttet over kommune- og fylkesgrenser, ofte for lange perioder. At denne praksisen har

hatt uheldig utslag og blant annet ført til feilplasseringer (Benedictow, 1986) er kjent, og hverken institusjonsbehandling (Kjellsberg & Dahl, 1998) eller fengsling (NOU, 2008) ser ut til å forebygge nye lovbrudd eller føre til andre positive langtidsresultater. De nye modell programmene bør derfor også vurderes i et familiepolitisk og oppvekstpolitisk perspektiv og som bidrag til utviklingen av en evidensbasert policy på atferdsfeltet. Det er uvanlig i Norge at det på så kort tid har blitt utviklet en omfattende forskningsbasert praksis rundt et prioritert problemområde. Både PMTO og MST legger, direkte og indirekte, vekt på prinsippet om familiebevaring (eng. family preservation) myndiggjøring av foreldre (eng. empowerment) og bygger på deres sterke sider slik at de kan hjelpe sine barn ut av problemene. Programmene ivaretar videre prinsippet om at problemer bør løses i miljøet hvor de oppstår og retten til å vokse opp på eget hjemsted bør så langt som mulig også gjelde for barn og unge med atferdsproblemer. Det er derfor ikke tilstrekkelig å spørre om programmene virker, men også om de har positive ringvirkninger på tjenestetilbudet og oppvekstmiljøet for barn og unge med alvorlige atferdsproblemer.

Det bør være en oppgave for videre implementeringsforskning å undersøke, ikke bare om en nasjonal implementering fører til at flere barn og unge får et behandlingstilbud, men om behandlingstilbudet blir mer tilgjengelig, og om den faglige behandlingskompetansen på atferdsfeltet øker. Hvis en vil undersøke om nasjonale prevalenstall for atferdsproblemer reduseres over tid, forutsetter det at en benytter reliable kriterier og har et egnet sammenligningsgrunnlag, nasjonalt eller i land som det er naturlig å sammenligne seg med. Norske studier har foreløpig ikke hatt fokus på kostnadene ved en bredt anlagt implementering av de nye programbaserte behandlingstilbudene. Det kan henge sammen med at "kostnad-nytte" analyser av norske behandlingstiltak foreløpig ikke har vært etterspurt, men kan være av interesse i framtidige intervensjonsstudier.

De familie- og nærmiljøbaserte programmene som er omtalt i denne avhandlingen har et verdigrunnlag og en tilnærming som synes å harmonere godt med norsk familiepolitikk på tvers av partigrenser og skiftende regjeringer. Selv om Norge har en sterk desentralisert struktur på barne- og ungdomstjenester, så bidrar prinsippet om likeverdige offentlige velferdstjenester i barnhager/skoler, helsestasjoner, pp-tjeneste, barnevern og psykisk helsevern til at nye intervensjoner i visse tilfeller kan implementeres over hele landet. Effektivitetsstudiene som inngår i denne avhandlingen hadde begrenset fokus på ringvirkningene som den nasjonale implementeringen og de positive behandlingsresultatene kan ha. Den økte innsatsen har utvilsomt ført til økt behandlingsskapasitet og høyere

kompetanse i arbeidet med alvorlige atferdsproblemer, men forhåpentlig også til et mer omfattende, variert og kvalitativt bedre tilbud til barn, ungdom og familier.

Referanser

- Achenbach, T.M. (1991). *Integrative guide for the 1991 CBCL/4-18, YSR and TRF profiles*. Burlington: University of Vermont, Department of Psychiatry.
- Addis, M.E., Wade, W.A., & Hatgis, C. (1999). Barriers to dissemination of evidence-based practices: Addressing practitioners' concerns about manualbased psychotherapies. *Clinical Psychology, Science and Practice*, 6, 430-441.
- Alexander, J. F., & Parsons, B. V. (1982). *Functional family therapy: Principles and procedures*. Carmel CA: Brooks & Cole.
- Alexander, J. F., Sexton, T. L., & Robbins, M. S. (2002). The development status of family therapy in family psychology intervention science. In Liddle, H.A. Santisteban, D.A., Levant, R., & Bray, J. (red.). *Family Psychology: Science-based interventions*. Washington DC: APA.
- Angell, B. Ø. (2003). Evidensbaserte programmer – kunnskapsformer og menneskesyn i sosialt arbeid. *Nordisk sosialt arbeid*, 23, 66-72.
- Aos,S., Lieb,R., Mayfield,J., Miller,M., & Pennuci,A. (2004). *Benefits and costs of prevention and early intervention programs for youth*. Olympia: Washington State Institute for Public Policy.
- Aos, S., Miller, M., & Drake, E. (2006). *Evidence-based adult correctional programs: What works and what does not?* Olympia: Washington State Institute for Public Policy.
- APA (2001). *Publication manual of the American Psychological Association*, 5th edn. Washington DC: American Psychological Association.
- APA (2002). Criteria for evaluating treatment guidelines. *American Psychologist*, 57, 1052-1059.
- APA (2006). Presidential task force on evidence-based practice. Evidence-based practice in psychology. *American Psychologist*, 61, 271-285.
- APA (2008). American Psychological Association Taks Force on Evidence-Based Practice for Children and Adolescents. *Disseminating evidence-based practice for children and adolescents: A systems approach to enhancing care*. Washington, DC: American Psychological Association.
- Armbruster, P. & Kazdin, A. E. (1994). Attrition in child psychotherapy. *Advances in Clinical Child Psychology*, 16, 81-108.

- Askeland, E., Christiansen, T., & Solholm, R. (2005). *PMTO håndbok*, Oslo: Atferdssenteret, Unirand.
- Asscher, J. J., Dekovic, M., Van der Laan, P.H., Prins, P. J.M., & Van Arum, S. (2007). Implementing randomized experiments in criminal justice settings: An evaluation of multi-systemic therapy in the Netherlands. *Journal of Experimental Criminology*, 3, 113-129.
- Bank, L., Marlowe, J. H., Reid, J. B., Patterson, G. R., & Weinrott, M. R. (1991). A comparative evaluation of parent-training evaluations for families of chronic delinquents. *Journal of Abnormal Child Psychology*, 19, 15-33.
- Baron, R. B., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51, 1173-1182.
- Beauchamp, M. H., & Anderson, V. (2010). SOCIAL: An integrative framework for the development of social skills. *Psychological Bulletin*, 136, 39-64.
- Beck, J. S. (1995). *Cognitive therapy: Basics and beyond*. Guilford Press, New York.
- Becker, W. C., Madsen, C. H., Arnold, R., & Thomas, D. R. (1967). The contingent use of teacher attention and praise in reducing classroom behavior problems. *Journal of Special Education*, 1, 287-307.
- Bendixen, M. & Olweus, D. (1999). Measurement of antisocial behavior in early adolescence and adolescence: Psychometric properties and substantive findings. *Criminal Behavior and Mental-Health*, 9, 323-354.
- Benedictow, M. (1996). *Ungdom i særtiltak*. Oslo: Universitetet i Oslo, Institutt for medisinsk atferdsfag.
- Bergman, Å., & Lundström, T. (2007). Att studera rörliga mål – om villkoren för evidens och kunskapsproduktion i socialt arbete. *Socionomens forsknings supplement*, 21, 4-16.
- Bernal, M. E., Klennert, M. D., & Schultz, L.A. (1980). Outcome evaluation of behavioral parent training and client-centered parent counseling for children with conduct problems. *Journal of Applied Behavior Analysis*, 13, 677-691.
- Best, A., Terpstra, J., Moor, G., Riley, B., Norman, C.D., & Glasgow, R. (2008). *Building knowledge integration systems for evidence-informed decisions*. Paper presented to the Organizational Behaviour in Health Care conference, Sydney, Australia, March 26th.
- Beutler, L. E., Bongar, B., & Shurkin, J. N. (1998). *A consumers guide to psychotherapy*. New York: Oxford University Press.

Beutler, L. E., Malik, M. L., Alimohamed, S. Harwood, T. M., Talebi, H., Noble, S. mfl. (2004). Therapist variables. I: Lambert, M. J. (red.) *Bergin and Garfield's handbook of psychotherapy and behavior change*. New York: Wiley.

Bickman, L. (1996). A continuum of care: More is not always better. *American Psychologist*, 51, 689-701.

Bickman, L., Guthrie, P.R., Foster, E.M., Lambert, E.M., Summerfelt, W.T., Breda, C.S., & Heflinger, C. (1995). *Evaluating managed mental health services: The Fort Bragg experiment*. New York: Plenum Press.

Bickman, L., Noser, K., & Summerfelt, W.T. (1999). Long-term effects of a system of care on children and adolescents. *Journal of Behavioral Health Services and Research*, 26, 185-202.

Biglan, T., & Ogden, T. (2008). The evolution of the evidence based movement. *European Journal of Behavior Analysis*, 9, 81-95.

Biglan, A., & Taylor, T. K. (2000). Increasing the use of science to improve child-rearing. *The Journal of Primary Prevention*, 21, 207-226.

Blow, A. J., Sprenkle, D. H., & Davis, S. D. (2007). Is who delivers the treatment more important than the treatment itself? The role of the therapist in commom factors. *Journal of Marital and Family Therapy*, 33, 298-317.

Bongers, I.L., Koot, H.M., van der Ende, J., & Verhulst, F.C. (2004). Developmental trajectories of externalizing behaviors in childhood and adolescence. *Child Development*, 75, 1523-1537.

Bordin, E. S. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, Research and Practice*, 16, 252-260.

Borduin, C. M., Mann, B. J., Cone, L. T., Henggeler, S. W., Fucci, B. R., Blaske, D. M., & Williams, R. A. (1995). Multisystemic treatment of serious juvenile offenders: Long-term prevention of criminality and violence. *Journal of Consulting and Clinical Psychology*, 63, 569-578.

Brestan, E.V., & Eyberg, S.M. (1998). Effective psychosocial treatments of conduct-disordered children and adolescents: 29 years, 82 studies, and 5,272 kids. *Journal of Clinical Child Psychology*, 27, 180-189.

Brestan, E.V., Eyberg, S. M., Boggs, S. R., & Algina, J. (1997). Parent-Child Interaction Therapy: Parents' perceptions of untreated siblings. *Child and Family Behavior Therapy*, 19, 13-28.

Brinkmeyer, M. Y., & Eyberg, S. M. (2003). Parent-child interaction therapy for oppositional children. I: Kazdin, A. E. & Weisz, J. R. (red.). *Evidence-based psychotherapies for children and adolescents*. New York: Guilford Press.

Bronfenbrenner, U. (1979). *The ecology of human development: Experiment by nature and design*. Cambridge MA: Harvard University Press.

Bullock, R. (1992). A framework for making international comparisons of services for adolescents. I: Bullock, R. (red.) *Problem adolescents. An international view*. London: Whiting and Birch Ltd.

Burns, B. (2003). Children and evidence-based practice. *Psychiatric Clinics of North America*, 26, 955-970.

Carpelan, K.S., Andershed, A-K., Andershed, H., Brännström, L., Nyström, M., & Ahlgren, T. (2008). *Insatser för unga lagöverträdare. En systematisk sammanställning av översikter om effekter på återfall i kriminalitet*. Rapport, Stockholm: Institutet för utveckling av metoder i socialt arbete (IMS).

Carroll, K.M., & Nuro, K.F. (2002). One size can't fit all: A stage model for psychotherapy manual development. *Clinical Psychology: Science and Practice*, 9, 396-406.

Carroll, K.M., & Rounsaville, B.J. (2008). Efficacy and effectiveness in developing treatment manuals. In: Nezu, A.M., & Nezu, C.M. (red.) *Evidence-based outcome research. A practical guide to conducting randomized controlled trials for psychosocial interventions*. Oxford: Oxford University Press.

Casey, R. J., & Berman, J. S. (1985). The outcome of psychotherapy with children. *Psychological Bulletin*, 98, 388-400.

Castonguay, L. G., Goldfried, M. R., Wiser, S., Raue, P. J., & Hayes, A. M. (1996). Predicting the effect of cognitive therapy for depression: A study of unique and common factors. *Journal of Consulting and Clinical Psychology*, 64, 497-504.

Chamberlain, P. (1998). *Family connections: A treatment foster care model for adolescence with delinquency*. Eugene OR: Northwest Media.

Chambless, D.L., & Hollon, S.D. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology*, 66, 7-18.

Chambless, D.L., & Ollendick, T.H. (2001). Empirically supported psychological interventions: Controversies and evidence. *Annual Review of Psychology*, 52, 685-716.

Christensen, B., & Mauseth, T. (2007). Multisystemisk terapi: familie- og nærmiljøbasert behandling av ungdom med alvorlige atferdsvansker. *Tidsskrift for Norsk Psykologforening*, 44, 1095-1106.

- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*, 2nd edn. Hillsdale NJ, Lawrence Erlbaum.
- Cohen, J. (1992a). Statistical power analysis. *Current Directions in Psychological Science*, 1, 98-101.
- Cohen, J. (1992b). A power primer. *Psychological Bulletin*, 112, 155-159.
- Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist*, 49, 997-1003.
- Cook, T.D. (2003). Why have educational evaluators chosen not to do randomized experiments? *The ANNALS of the American Academy of Political and Social Science*, 589, 114-149.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-Experimentation design and analysis issues for field settings*. Skokie IL: Rand McNally.
- Cooper, P., C. J. Smith og G. Upton (1994). *Emotional behavioural difficulties: Theory to practice*. London: Routledge.
- Crick, N., & Dodge, K. (1994). A review and reformulation of social information-processing mechanisms in children's social adjustment. *Psychological Bulletin*, 115, 74-101.
- Crits-Christoph, P., Baranackie, K., Kurcias, J.S., Beck, A.T., Carroll, K. Perry, K., Luborsky, L., McLellan, A.T., Woody, G.E., Thompson, L., Gallagher, D., & Zitrin, C. (1991). Meta-analysis of therapist effects in psychotherapy outcome studies. *Psychotherapy Research*, 1, 81-91.
- Crits-Cristoph, P., Wilson, G. T., & Hollon, S. D. (2005). Empirically supported psychotherapies: Comment on Westen, Novotny, and Thompson-Brenner (2004). *Psychological Bulletin*, 131, 412-417.
- Cunningham, A. (2002, juni). *Lessons learned from a randomized study of Multisystemic Therapy in Canada*. PRAXIS: Research from the Centre for Children and Families in the Justice System. London/Canada, Centre for Children and Families in the Justice System (www.lfcc.on.ca).
- Curtis, N.M., Ronan, K.R., & Borduin, C.M. (2004). Multisystemic treatment: A meta-analysis of outcome studies. *Journal of Family Psychology*, 18: 411-419.
- Curtis, N. M., Ronan, K. R., Heiblum, N., & Kylie, C. (2009). Dissemination and effectiveness of multisystemic treatment in New Zealand: A benchmarking study. *Journal of Family Psychology*, 23, 119-129.
- Davidson, K. W., Goldstein, M., Kaplan, R. M., Kaufmann, P. G., Knatterud, G. L., Orleans, C. T. mfl. (2003). Evidence-based behavioral medicine: What is it, and how do we achieve it? *Annals of Behavioral Medicine*, 26, 161-171.

Davies, P. (2004). "Is evidence-based government possible?" Jerry Lee lecture 2004. The 4th annual Campbell Collaboration colloquium, Washington DC.

DeGarmo, D. S., Patterson, G. R., & Forgatch, M. S. (2004). How do outcomes in a specified parent training intervention maintain or wane over time? *Prevention Science*, 5, 73-89.

Dishion, J.T., McCord, J., & Poulin, F. (1999). When interventions harm: Peer groups and problem behavior. *American Psychologist*, 54, 1-10.

Dishion, T.J., & Patterson, G.R. (2006). The development and ecology of antisocial behavior in children and adolescents. I: Cicchetti, D. & Cohen, D.J. (red.), *Developmental psychopathology. Vol 3: Risk disorder and adaptation*, New York: Wiley.

Dodge, K. A., Dishion, T. J., & Lansford, J. E. (2006). Deviant peer influences in intervention and public policy for youth. *Social Policy Report*, Vol XX, 1-19.

Drake, E. K., Aos, S., & Miller, M. G. (2009). Evidence-based public policy options to reduce crime and criminal justice costs: Implications in Washington State. *Victims and Offenders*, 4, 170-196.

Duncan, B. L., & Miller, S. D. (2000). *The heroic client: Doing client directed, outcome-informed therapy*. San Francisco: Jossey-Bass.

Duncan, B. L., & Miller, S. D. (2005). Treatment manuals do not improve outcomes. I: Norcross, J, Levant, R., & Beutler, L. (red.). *Evidence-based practices in mental health: Debate and dialogue on the fundamental questions*. Washington D.C. : American Psychological Association Press.

Durlak, J.A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology*, 41, 327-350.

Econ-rapport nr 2009-117 (2009). *Hvorfor blir det flere unge uføre?* Utarbeidet for Arbeids- og inkluderingsdepartementet, Oslo: Econ, Pöyry AS.

Eddy, J. M., & Chamberlain, P. (2000). Family management and deviant peer association as mediators of the impact of treatment condition on youth antisocial behavior. *Journal of Consulting and Clinical Psychology*, 68, 857-863.

Ekeland, T. J. (1999). Evidensbasert behandling: kvalitetssikring eller instrumentalistisk mistak? *Tidsskrift for Norsk Psykologforening*, 36, 1036-1047.

Elkin, I., Gibbons, R.D., Shea, M.T., Sotsky, S.M., Watkins, J.T., Pilkonis, P.A. (1995). Initial severity and differential treatment outcome in the National Institute of Health

Treatment of Depression Collaborative Research Program. *Journal of Consulting and Clinical Psychology*, 63, 841-847.

Elliott, D. S. (1998). *Blueprints for violence prevention*. Boulder, CO: University of Colorado, Center for the Study and prevention of Violence. Blueprints Publications.

Elliott, D. S., Ageton, S. S., Huizinga, D., Knowles, B. A., & Canter, R. J. (1983). *The prevalence and incidence of delinquent behaviour: 1976-80. (Report of the National Youth Survey, Project Rept. #26)*. Boulder, CO: Behavioural Research Institute.

Elliott, D. S., Huizinga, D., & Ageton, S. S. (1985). *Explaining delinquency and drug use*. Beverly Hills CA: Sage.

Elliott, D.S., & Mihalic, S. (2004). Issues in disseminating and replicating effective prevention programs. *Prevention Science*, 5, 47-53.

Eme, R. F. (1992). Selective female affliction in the developmental disorders of childhood: A literature review. *Journal of Clinical Child Psychology*, 21, 354-364.

Emmelkamp, P.M.G., Bouman, T., Blaauw, E. (1994). Individualized vs. Standardized therapy: a comparative evaluation with obsessive-compulsive patients. *Clinical Psychology and Psychotherapy*, 1, 95-100.

Eyberg, S. H., Nelson, M. M., & Boggs, S. R. (2008). Evidence-based psychosocial treatments for children and adolescents with disruptive behavior. *Journal of Clinical Child and Adolescent Psychology*, 37, 215-237.

Farrington, D.P., & Welsh, B.C. (2003). Family-based prevention of offending: A meta-analysis. *Australian and New Zealand Journal of Criminology*, 36, 127-151.

Ferrer-Wreder, L., Stattin, H., Lorente, C.C., Tubman, J.G., & Adamson, L. (2004). *Successful prevention and youth development programs across borders*. New York: Kluwer Academic.

Fixsen, D. L., Naoom, S. F., Blasé, K. A., Friedman, R. M., & Wallace, F. (2005). *Implementation research: A synthesis of the literature*. Available from <http://www.fpg.unc.edu/~nirn/resources/publications/Monograph/>

Fixsen, D.L., Blase, K. A., Naoom, S. F., & Wallace, F. (2009). Core implementation Components. *Research on Social Work Practice*, 19, 531-540.

Flay, B.R., Biglan, A., Boruch, R.F., Castro, F.G., Gottfredsson, D., Kellam, S., Moscicki, E.K., Schinke, S., Valentine, J.C., & Ji, P. (2005). Standards of evidence: Criteria for efficacy, effectiveness and dissemination. *Prevention Science*, 6, 151-175.

Forehand, R. L. & Long, N. (1991) Prevention of aggression and other behavior problems in the early adolescent years. I: Pepler, D.J. & Rubin, K.H. (red.) *The development and treatment of childhood aggression*, Hillsdale, NJ: Erlbaum.

Forgatch, M. S. (1991). The clinical science vortex: A developing theory of antisocial behavior. In Pepler, D. J. & Rubin, K.H. (red..) *The development and treatment of childhood aggression*, pp. 291-313. Hillsdale: Lawrence Erlbaum Associates.

Forgatch, M. S. (1994). *Parenting through change: A programmed intervention curriculum for groups of single mothers*. Eugene: Oregon Social Learning Center.

Forgatch, M. S., & DeGarmo, D. S. (1999). Parenting through change: An effective prevention program for single mothers. *Journal of Consulting and Clinical Psychology*, 67, 711-724.

Forgatch, M. S., & DeGarmo, D. S. (2002). Extending and testing the social interaction learning model with divorce samples. In Reid, J.B., Patterson, G.R., & Snyder, J. (red.), *Antisocial behavior in children and adolescents: A developmental analysis and model for intervention*. Washington DC: American Psychological Association.

Forgatch, M. S., DeGarmo, D. S., & Beldavs, Z. (2005). An efficacious theory-based intervention for stepfamilies. *Behavior Therapy*, 36, 357-365.

Forgatch, M. S., Patterson, G. R., DeGarmo, D. S., & Beldavs, Z. G. (2009). Testing the Oregon delinquency model with 9-year follow-up of the Oregon divorce study. *Development and psychopathology*, 21, 637-660.

Fossestøl, K., & Enehaug, H. (2008). *Konservative byråkrater eller avantgardistiske byråkrater (Conservative practitioners or avantgarde bureaucrats)*. Rapport. Oslo: Arbeidsforskningsinstituttet.

Frank, J. D., & Frank, J. B. (1991). *Persuasion and healing: A comparative study of psychotherapy (3rd edn.)*. Baltimore: John Hopkins University Press.

Frick, P. J. (1998). Conduct disorders. I: Ollendick, T.H. and Hersen, M. (red). *Handbook of child psychopathology*, 3rd ed. New York, NY: Plenum Press.

Gibaldi, M. & Sullivan, S. (1997). Intention-to-treat analysis in randomized trials: Who gets counted? *Journal of Clinical Pharmacology*, 37, 667-672.

Gibbs, L., & Gambrill, E. (2002). Evidence-based practice: Counterarguments to objections. *Research on Social Work Practice*, 12, 452-476.

Glasgow, R. E., (2009). Critical measurement issues in translational research. *Research on Social Work Practice*, 19, 560-568.

Glasgow, R. E., Vogt, T. M., & Boles, S. M. (1999). Evaluating the public health impact of health promotion interventions: the RE-AIM framework. *American Journal of Public Health*, 89, 1322-1327.

Glasgow, R. E., Lichtenstein, E., Marcus, A. C. (2003). Why don't we see more translation of health promotion research to practice? Rethinking the efficacy-to-effectiveness transition. *American Journal of Public Health*, 93, 1261-1267.

Greenberg, M.T., Domitrovich, C.E., Graczyk, & Zins, J.E. (2005). *The study of implementation in school-based preventive interventions: Theory, research and practice (Volume 3)*. Rockville, MD: Center for Mental Health Services, Substance Abuse and Mental Health Services Administration.

Gresham, F. M., Elliott, S. N. (1990). *Social Skills Rating System*. Circle Pines, MN: American Guidance Service.

Haley, J. (1976). *Problem solving therapy*. San Francisco: Jossey-Bass.

Hanna, F. J. (1996). Precursors of change: Pivotal points of involvement and resistance in psychotherapy. *Journal of Psychotherapy Integration*, 6, 227-264.

Hansen, N.B., Lambert, M.J., & Foreman, E.M. (2002). The psychotherapy dose-response effect and its implications for treatment delivery services. *Clinical Psychology: Science and Practice*, 9, 329-343.

Hatland, A., & Bay, A-H. (2009). Forskningens rolle i velferdspolitikken: instrumentell eller strategisk? I: B.J. Nuland, B.S. Tranøy og J. Christensen (red.). *Hjernen er alene. Institusjonalisering, kvalitet og relevans i norsk velferdsforskning*. Oslo: Universitetsforlaget.

Haugsgjerd, S., Jensen, P., Karlsson, B., & Løkke, J.A. (2009). *Perspektiver på psykisk lidelse – å forstå, beskrive og behandle*. Oslo: Gyldendal Akademisk.

Heiervang, E., Stormark, K. M., Lundervold, A. J., Heiman, M., Goodman, R., Posserud, M. B., Ullebo, A. K., Plessen, K.J., Bjelland, I., Lie, S. A., & Gilleberg, C. (2007). Psychiatric disorders in Norwegian 8-to-10-year-olds: An epidemiological survey of prevalence, risk factors, and service use. *Journal of the American Academy of Child and Adolescent Psychiatry*, 46, 438-447.

Henggeler, S. W. (1991). A consensus: Conclusions of the APA task force report on innovative models of mental health services for children, adolescents, and their families. *Journal of Clinical Child Psychology*, 23, 3-6.

Henggeler, S. W. (1997). "The development of effective drug abuse services for youth." I : Egertson, J. A., Fox, D.M. 6 Leshner, A.I. (red.): *Treating drug abusers effectively*. New York: Blackwell.

Henggeler, S.W. (2004). Decreasing effect sizes for effectiveness studies – Implications for the transport of evidence-based treatments: Comment on Curtis, Ronan and Borduin *Journal of Family Psychology*, 18, 420-423..

Henggeler S. W. & Borduin, C. M. (1992). *Multisystemic therapy adherence scales*. Unpublished instrument. Department of Psychiatry and Behavioural Sciences, Medical University of South Carolina.

Henggeler, S. W., & Lee, T. (2003). Multisystemic treatment of serious clinical problems. I: A. E. Kazdin & J. R. Weisz (red.) *Evidence-based psychotherapies for children and adolescents*, New York: Guilford.

Henggeler, S. W., Melton, G. B., Brondino, M. J., Scherer, D. G., & Hanley, J. H. (1997). Multisystemic therapy with violent and chronic juvenile offenders and their families: The role of treatment fidelity in successful dissemination. *Journal of Consulting & Clinical Psychology*, 65, 821-833.

Henggeler, S. W., Melton, G. B., & Smith, L. A. (1992). Family preservation using multisystemic therapy: An effective alternative to incarcerating serious juvenile offenders. *Journal of Consulting & Clinical Psychology*, 60, 953-961.

Henggeler, S. W., Melton, G. B., Smith, L. A., Schoenwald, S. K., & Hanley, J. H. (1993). Family preservation using multisystemic treatment: Long-term follow-up to a clinical trial with serious juvenile offenders. *Journal of Child and Family Studies*, 2, 283-293.

Henggeler, S. W., Pickrel, S. G., & Brondino, M. J. (1999). Multisystemic treatment of substance abusing and dependent delinquents: Outcomes, treatment fidelity, and transportability. *Mental Health Services Research*, 1, 171-184.

Henggeler, S. W., Schoenwald, S. K., Borduin, C. M., Rowland, M. D., & Cunningham, P. B. (1998). *Multisystemic Treatment of Antisocial Behaviour in Children and Adolescents*. New York: The Guilford Press.

Henggeler, S. W., Schoenwald, S. K., Borduin, C. M., Rowland, M. D., & Cunningham, P. B. (2000). *Multisystemisk behandling av barn og unge med atferdsproblemer*. Oslo: Kommuneforlaget.

Henggeler, S. W., Schoenwald, S. K., Borduin, C. M., Rowland, M. D., & Cunningham, P. B. (2009). *Multisystemic Treatment of Antisocial Behaviour in Children and Adolescents*. 2nd edition. New York: The Guilford Press.

- Henggeler, S. W., Schoenwald, S. K., Swenson, C. C. & Borduin, C. M. (2006). Methodological critique and meta-analysis as a Trojan horse. Letter to the editor. *Children and Youth Services Review*, 28, 447-457.
- Horvarth, A. O., Symmonds, D. B. (1991). Relationship between working alliance and outcome in psychotherapy: a meta-analysis. *Journal of Counseling Psychology*, 38, 139-149.
- Hubble, M. A., Duncan, B. L., & Miller, S. D. (red.) (1999). *The heart and soul of change: What works in therapy?* Washington DC: American Psychological Association.
- Huey, S. J., Jr., Henggeler, S. W., Brondino, M. J., & Pickrel, S. G. (2000). Mechanisms of Change in Multisystemic Therapy: Reducing Delinquent Behavior through Therapist Adherence and Improved Family and Peer Functioning. *Journal of Consulting and Clinical Psychology*, 68, 451-467.
- ICD-10. (1999). *Psykiske lidelser og atferdsforstyrrelser. Kliniske beskrivelser og diagnostiske retningslinjer*. Statens Helsetilsyn. Oslo: Universitetsforlaget.
- Innstilling om *Verneskolenes funksjon og målsetting* (1968) Oslo: Sosialdepartementet.
- Jacobson, N. S., Schmalings, K. B., Holtzworth-Munroe, A., Katt, J. L., Wodd, L. F. mfl. (1989). Research structured vs. clinically flexible versions of social-learning based marital therapy. *Behavioral Research Therapy*, 27, 173-180.
- Jensen, P. S., Weersing, R., Hoagwood, K. E. , & Goldman, E. (2005). What is the evidence for evidence-based treatments? A hard look at our soft underbelly. *Mental Health Services Research*, 7, 53-74,
- Jergeby, U. (red.) (2008). *Evidensbaserad praktik i socialt arbete*. Stockholm: Gothia Förlag.
- Jergeby, U. & Sundell, K. (2008). Evidensbaserat socialt arbete. I: Jergeby, U. (red.). *Evidensbaserad praktik i socialt arbete*. Stockholm: Gothia Förlag.
- Kaufmann, A. H. (1988). *Antisocial atferd hos ungdom: En studie av psykologiske determinanter*. Oslo: Sigma Forlag A/S.
- Kaminski, J.W., Valle, L.A., Filene, J.H., & Boyle, C.L. (2008). A meta-analytic review of components associated with parent training program effectiveness. *Journal of Abnormal Child Psychology*, 36, 567-589.
- Kazdin, A. E. (1978). *History of behavior modification: Experimental foundations of contemporary research*. (xi ed.) Baltimore, MD, US: University Park Press.
- Kazdin, A. (1986). Comparative outcome studies of psychotherapy: Methodological issues and strategies. *Journal of Consulting and Clinical Psychology*, 62, 95-105.

Kazdin, A. E. (1997). Conduct disorder across the life-span. In Luthar, S. S., & Burack, J.A. (red.) *Developmental psychopathology. Perspectives on adjustment, risk, and disorder*. New York: Cambridge University Press.

Kazdin, A. (2008). Evidence-based treatment and practice. New opportunities to bridge clinical research and practice, enhance the knowledge base, and improve patient care. *American Psychologist*, 63, 146-159.

Kazdin, A. E., & Nock, M. K. (2003). Delineating mechanisms of change in child and adolescent therapy: methodological issues and research recommendations. *Journal of Child Psychology and Psychiatry*, 44, 1116-1129.

Kellam, S.G., & Langevin, D.J. (2003). A framework for understanding “evidence” in prevention research and programs. *Prevention Science*, 4, 137-153.

Kendall, P. C., Chu, B., Gifford, A., Hayes, C., Nauta, M. (1998). Breathing life into a manual: flexibility and creativity with manual-based treatments. *Cognitive Behavioral Practitioner*, 5, 177-198.

Kjellsberg, E. & Dahl, A. A. (1998). High delinquency, disability and mortality - a register study of former adolescent psychiatric in-patients. *Acta Psychiatrica Scandinavia*, 98, 34-40.

Kjøbli, J., & Ogden, T. (2009) Gender differences in intake characteristics and behavior change among children in families receiving parent management training. *Children and Youth Services Review*, 31, 823-830.

Klein, D.F. (1996). Preventing hung juries about therapy studies. *Journal of Consulting and Clinical Psychology*, 64, 81-87.

Knutson, N. M., Forgatch, M. S., & Rains, L. A. (2003). *Fidelity of implementation rating system (FIMP): The training manual for PMTO*. Eugene, OR: Social Learning Center.

Krupnik, J. L., Sotsky, S. M., Simmens, S., Moyer, I., Elkin, I. mfl. (1996). The role of the therapeutic alliance in psychotherapy and pharmacotherapy outcome: findings in the NIMH treatment of Depression Collaborative Research Program, *Journal of Consulting and Clinical Psychology*, 64, 532-539.

LaGreca, A.M., Silverman, W. K., & Lochman, J. E. (2009). Moving beyond efficacy and effectiveness in child and adolescent intervention research. *Journal of Consulting and Clinical Psychology*, 77, 373-382.

Lambert, M. J. (1992). Psychotherapy outcome research: Implications for integrative and eclectic therapists. I: Norcross, J.C., & Goldfried, M.R. (red.) *Handbook of therapy integration*. New York: Basic Books.

Lambert, M. J., & Ogles, B. M. (2004). The efficacy and effectiveness of psychotherapy. I: Lambert, M.J. (red.), *Bergin and Garfield's handbook of psychotherapy and behavior change (5th edn)*, New York: Wiley.

Larsson, B., Fossum, S., Clifford, G., Drugli, M. B., Handegård, B. H. & Mørch, W-T. (2009). Treatment of oppositional defiant and conduct problems in young Norwegian children: Results of a randomized controlled replication study. *European Child and Adolescent Psychiatry*, 18, 42-52.

Lavik, N. J. (1976). *Ungdoms mentale helse*. Oslo: Universitetsforlaget.

Lescheid, A., & Cunningham, A. (2002 February). *Seeking effective interventions for serious young offenders: Interim results of a four-year randomized study of multisystemic therapy in Ontario, Canada*. London, Centre for Child and Families in the Justice system.

Lipsey, M. (1995). Her sitert fra A. Petrosino (2004). *Presentation prepared for Social methods seminar, 9 december, 2004, Centre for evaluation of social services, National Board of health and welfare, Stockholm*

Lipsey, M.W. (1999). Can rehabilitative programs reduce the recidivism of juvenile offenders? An inquiry into the effectiveness of practical programs. *Virginia Journal of Social Policy & the Law*, 6, 611-641.

Lipsey, M.W. (2000). Statistical conclusion validity for intervention research: A significant ($p < .05$) problem. I: Bickman, L. (red.) *Validity and social experimentation: Donald Campbell's legacy*. Thousand Oaks: CA: Sage.

Lipsey, M.W. & Cordray, D. S. (2000). Evaluation methods for social interventions. *Annual Review of Psychology*, 51, 345-375.

Lipsey, M.W., & Cullen, F.T. (2007). The effectiveness of correctional rehabilitation: A review of systematic reviews. *The Annual Review of Law and Social Science*. 3, 297-320.

Littell, J.H. (2005a). Lessons learned from a systematic review of multisystemic therapy. *Children and Youth Services Review*, 27, 445-463.

Littell, J.H. (2005b). The case for Multisystemic Therapy: Evidence or orthodoxy? *Children and Youth Services Review*, 28, 458-472.

Littell, J.H., Popa, M. & Forsythe, B. (2005). *Multisystemic therapy for social, emotional and behavioural problems in youth aged 10-17*. Rapport 05:01, Nordic Campbell Center www.nc2.net

Loeber, R. (1990). Development and risk factors of juvenile antisocial behavior and delinquency. *Clinical Psychology Review*, 10, 1-41.

Loeber, R., & Farrington, D. P. (1998). *Serious & violent juvenile offenders: Risk factors and successful interventions*. Thousand Oaks: Sage Publications.

Loeber, R., & Hay, D.F. (1994). Developmental approaches to aggression and conduct problems. I: Rutter, M. & Hay, D.F. (red.), *Development through life: A handbook for clinicians*. Oxford: Blackwell Scientific.

Luborsky, L., Singer, B., & Luborsky, L. (1975). Comparative studies of psychotherapies: Is it true that "Everyone has won and all must have prizes?" *Archives of General Psychiatry*, 32, 995-1008.

Luborsky, L., Diguer, L., Seligman, D.A., Rosenthal, R., Krause, E.D., & Johnson, S., mfl. (1999). The researcher's own therapy allegiances: A "wild card" in comparisons of treatment efficacy. *Clinical Psychology: Science and Practice*, 4, 53-65.

Lubrecht, J. (1992). Family Satisfaction Survey. I: Kutash, K. & Rivera, T.R. *Measures of Satisfaction with Child Mental Health Services*. Unpublished manuscript, Florida Mental Health Institute, University of South Florida.

Lundahl, B., Risser, H.J., & Lovejoy, M.C. (2006). A meta-analysis of parent training: Moderators and follow-up effects. *Clinical Psychology Review*, 26, 86-104.

Löfholm, C.A., Olsson, T., Sundell, K., & Hansson, K. (2009). Multisystemic therapy with conduct-disordered young people: stability of treatment outcomes two years after intake. *Evidence and Policy*, 5, 373-397.

Lösel, F. (1993). Evaluation of psychosocial interventions in prison and other penal contexts. I *Psychosocial interventions in the criminal justice system*. Proceedings. Report presented to the 20th Criminological Research Conference. Council of Europe. European Committee on Crime Problems. Criminological research, Vol XXXI.

Madsen, C.H., Becker, W.C., & Thomas, D.R. (1968). Rules, praise, and ignoring: Elements of elementary classroom control. *Journal of Applied Behavior Analysis*, 1, 139-150.

Mann, B. J. Borduin, C. M., Henggler, S. W. & Blaske, (1990). An investigation of systemic conceptualizations of parent-child coalition and symptom change. *Journal of Consulting and Clinical Psychology*, 58: 336-344.

Martinez, C. R. jr., & Forgatch, M. (2001). Preventing problems with boys' noncompliance: Effects of a parent training intervention for divorcing mothers. *Journal of Consulting and Clinical Psychology*, 69, 416-428.

Martinson, R. (1974). What works? Questions and answers about prison reform. *The Public Interest*, 35, 22-54.

Mathiesen, K. S., & Sanson, A. (2000). Dimensions of early childhood problems: Stability and predictors of change from 18 to 30 months. *Journal of Abnormal Child Psychology*, 28, 15-31.

McCart, M. R., Priester, P. E., Davies, W. H., & Azen, R. (2006). Differential effectiveness of behavioral parent-training and cognitive-behavioral therapy for antisocial youth: A meta-analysis. *Journal of Abnormal Child Psychology*, 34, 527-543.

Meehl, P.E. (1954). *Clinical vs. statistical prediction*. Minneapolis: University of Minnesota Press.

Meichenbaum, D. (1993). Changing conceptions of cognitive behavior modification: Retrospect and prospect. *Journal of Consulting and Clinical Psychology*, 61, 202-204.

Mihalic, S., Fagan, A., Irwin, K., Ballard, D., & Elliott, D. (2004). *Blueprints for violence prevention*. Available from www.ojp.usdoj.gov/ojjdp

Mihalic, S.F., & Irwin, K. (2003). Blueprints for violence prevention: From research to real world settings. Factors influencing the successful replication of model programs. *Youth violence and juvenile justice*, 1, 1-23.

Miller, G. E., & Prinz, R. J. (1990). Enhancement of social learning family interventions for childhood conduct disorder. *Psychological Bulletin*, 108, 291-307.

Minuchin, S. (1974). *Families and family therapy*. Cambridge, MA: Harvard University Press.

Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review*, 100, 674-701.

Moffitt, T. E., & Caspi, A. (2001). Childhood predictors differentiate life-course persistent and adolescence-limited antisocial pathways, among males and females. *Development and Psychopathology*, 13, 355-375.

Moffitt, T.E., Caspi, A., Rutter, M., & Silva, P.A. (2001). *Sex differences in antisocial behaviour. Conduct disorder, delinquency and violence in the Dunedin longitudinal study*. Cambridge: Cambridge University Press.

Moher, D., Dulberg, C. S., & Wells, G. A. (1994). Statistical power, sample size and their reporting in randomized controlled trials. *Journal of American Medical Association*, 272, 122-124.

Moher, D., Hopewell, S., Schulz, K. F., Montori, V., Gøtzsche, P. C., Devereaux, P. J., Elbourne, D., Egger, M., & Altman, D. G. (2010). CONSORT 2010 Explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *Research Methods & Reporting*. BMJ 2010; 340: c869 (doi: 10.1136/bmj.c869).

Moher, D., Schulz, K.F., & Altman, D.G. (2001). The CONSORT statement: revised recommendations for improving the quality of reports of parallel group randomized trials. *Lancet*, 357, 1191-1194.

Munger, R. L. (1993). *Changing children's behavior quickly*. Madison, IL: Lanham.

Nathan, P.E., & Gorman, J.M. (red.) (2007). *A guide to treatment that work*. 3rd edn. Oxford: Oxford University Press.

Nezu, A.M. & Nezu, C.M. (red.). *Evidence-based outcome research. A practical guide to conducting randomized controlled trials for psychosocial interventions*. Oxford: Oxford University press.

Nock, M. K., Janis, I. B, & Wedig, M. M. (2008). Research designs. I: Nezu, A.M. & Nezu, C.M. (red.). *Evidence-based outcome research. A practical guide to conducting randomized controlled trials for psychosocial interventions*. Oxford: Oxford University press.

Norcross, J.C. (red.) (2002). *Psychotherapy relations that work: Therapist contributions and responsiveness to patient needs*. New York: Oxford University Press.

Norges offentlige utredninger (NOU 2008:15) *Barn og straff – utviklingsstøtte og kontroll*. Oslo: Justisdepartementet.

Norges offentlige utredninger (NOU 2009:22). *Det du gjør, gjør det helt*. Oslo: Barne- og likestillingsdepartementet.

Nutley, S., Morton, S., Jung, T., & Boaz, A. (2010). Editorial: Evidence and policy in Northwestern Europe. *Evidence & Policy*, 6.

Nutley, S., Walter, I., Davies, H.T.O. (2007). *Using Evidence. How research can inform public services*. Bristol: The Policy press.

Nøvik, T. S. (1999). *Validity and use of the Child Behaviour Checklist in Norwegian children and adolescents. An epidemiological and clinical study*. Doctoral dissertation, Oslo: Centre for child and adolescent psychiatry, Department group of psychiatry, University of Oslo.

Ogden, T. (1998). Multisystemisk terapi. *Spesialpedagogikk*, 5, 16-24.

Ogden, T. (1999). Antisocial atferd og barneoppdragelse: "Parent management training" som foreldreopplæring. *Spesialpedagogikk*, 6, 3-17.

Ogden, T. (2002). Kartlegging og utredning av ungdom med alvorlige atferdsproblemer. I: Rønnestad, M.H. & Lippe, A. (red.) *Det kliniske intervjuet*. Oslo: Gyldendal Akademisk.

Ogden, T. (2003). The Validity of Teacher Ratings of Adolescents' Social Skills. *Scandinavian Journal of Educational Research*, 47, 63–76.

Ogden, T., & Amlund-Hagen, K. (2006). Multisystemic treatment of serious behaviour problems in youth: Sustainability of effectiveness two years after intake. *Child and Adolescent Mental Health*, 11, 142-149.

Ogden, T., & Amlund-Hagen, K. (2006a). Virker MST? Kommentarer til en systematisk forskningsoversikt og meta-analyse av MST. *Nordisk sosialt arbeid*, 26, 222-233.

Ogden, T. & Amlund-Hagen, K. (2008). Treatment effectiveness of Parent Management Training in Norway: A randomized controlled trial of children with conduct problems. *Journal of Consulting and Clinical Psychology*, 76, 607-621.

Ogden, T. & Amlund-Hagen, K. (2009). What works for whom? Gender differences in intake characteristics and treatment benefit of Multisystemic Therapy. *Journal of Adolescence*, 32, 1425-1435.

Ogden, T., Amlund-Hagen, K., Askeland, E., & Christensen, B. (2009). Implementing and evaluating evidence-based treatments of conduct problems in children and youth in Norway. *Research on Social Work Practice*, 19, 582-591.

Ogden, T., Christensen, B., Sheidow, A., & Holth, P. (2008). Bridging the gap between science and practice: The effective nationwide transport of MST programs in Norway. *Journal of Children and Adolescent Substance Abuse*, 17, 93-109.

Ogden, T., Forgatch, M., Askeland, E., Patterson, G.R., & Bullock, B. (2005). Implementation of Parent Management Training at the national level: The case of Norway. *Journal of Social Work Practice*, 19, 317-329.

Ogden, T. & Halliday-Boykins, C.A. (2004). Multisystemic treatment of antisocial adolescents in Norway: Replication of clinical outcomes outside of the US. *Child and Adolescent Mental Health*, 9, 76-82.

Ogden, T., Kärki, F.U, & Teigen, K. S. (2010). Linking research, practice and policy in the welfare services and education in Norway, *Evidence and Policy*. (under trykking).

Olds, D. L., Robinson, J., O'Brian, R. mfl. (2002). Home visits by paraprofessionals and by nurses: A randomized controlled trial. *Pediatrics*, 110, 486-496.

Olson, D., Portner, J., & Lavee, Y. (1985). *FACES III*. University of Minnesota: Department of Family Social Service.

- Olsson, T. (2008). *Crossing the quality chasm?* Lund: Lund dissertations in social work.
- Olweus, D. (1992). *Mobbing i skolen: Hva vet vi og hva kan vi gjøre ?* Oslo: Universitetsforlaget.
- Onken, L.S., Blaine, J.D., & Battjes, R. (1997). Behavioral therapy research: A conceptualization of a process. In Henggeler, S.W. & Amentos, R. (red.), *Innovative approaches for difficult to treat populations*. Washington, DC: American Psychiatric Press.
- Otnes, A., & Prydz, M. B. (2009). *Multisystemisk terapi (MST) og atferdsvansker hos ungdom: Er alder relatert til endring?* Hovedoppgave, Psykologisk Institutt, Oslo: Universitetet i Oslo.
- Patterson, G. R. (1982). *Coercive family process*. Eugene, OR: Castalia.
- Patterson, G. R. (1986). Performance models for antisocial boys. *American Psychologist*, 41, 432-444.
- Patterson, G. R. (2002). The early development of the coercive family process. I: Reid, J.B., Patterson, G. R., & Snyder, J.J. (red.): *Antisocial behavior in children and adolescents: A developmental analysis and a model for intervention*. Washington DC: American Psychological Association.
- Patterson, G. R. (2005). The next generation of PMTO models. *The Behavior Therapist*, 35, 27-33.
- Patterson, G.R., Dishion, T. J., & Chamberlain, P. (1993) Outcomes and methodological issues relating to treatment of antisocial children. I: Giles, T.R. (red.) *Handbook of effective psychotherapy*, 43-88. New York: Plenum Press.
- Patterson, G. R., Chamberlain, P., & Reid, J. B. (1982). A comparative evaluation of a parent-training program. *Behavior Therapy*, 13, 638-650.
- Patterson, G. R. & Forgatch, M. S. (1995). Predicting future clinical adjustment from treatment outcomes and process variables. *Psychological Assessment*, 7, 275-285.
- Patterson, G.R. & Forgatch, M. S. (2000). *Å leve sammen: Foreldre og tenåringer: Det grunnleggende*. Del 1. Oslo: PAX.
- Patterson, G. R., & Forgatch, M. S. (2005). *Parents and adolescents living together: Part 1: The basics*. Champaign, IL: Research Press.
- Patterson, G. R., Reid, J. B., Jones, R. R., & Conger, R. E. (1975). *A social learning approach to family intervention: Families with aggressive children. (Vol 1)*. Eugene OR: Castalia.

Patterson, G. R., Reid, J. B., & Dishion, T. J. (1992). *A social interactional approach: Vol. 4: Antisocial boys*. Eugene, OR: Castalia.

Pereplechikova, F., & Kazdin, A. E. (2005). Treatment integrity and therapeutic change: Issues and research recommendations. *Clinical Psychology: Science and practice*, 12, 365-383.

Pereplechikova, F., Hilt, M. L., Chereji, E., Kazdin, A. E. (2009). Barriers to implementing treatment integrity procedures: Survey of treatment outcome researchers. *Journal of Consulting and Clinical Psychology*, 77, 212-218.

Persons, J.B. (1991). Psychotherapy outcome studies do not accurately represent current models of psychotherapy: A proposed remedy. *American Psychologist*, 46, 99-106.

Petrosino, A. (2004). *The impact of program developers as evaluators: Evidence from prior and current offender treatment meta-analyses*. Presentation prepared for the Social methods seminar, 9. desember 2004, Stockholm, Centrum för Utvärdering av metoder i Socialt arbete (CUS).

Pettersson, T. (2010). *Återfall i brott bland ungdomar dömda till fängelse respektive slutet ungdomsvård*. Rapport 2/10, Stockholm: Statens Institutionsstyrelse.

Piquero, A., Farrington, D.P., Welsh, B., Tremblay, R., & Jennings, W. (2008). *Effects of early family/parent training programs on antisocial behavior and delinquency. A systematic review*. Rapport. Stockholm, Swedish National Council for Crime Prevention.

Price, R. H. (2002). *The four faces of community readiness: Social capital, problem awareness, social innovation and collective efficacy*. Manuscript. University of Michigan, Institute of social research.

Price, R.H., & Lorion, R. P. (1989). Prevention programming as organizational reinvention: From research to implementation. I: Shaffer,D., Philips, I., & Enzer, N.B. (red.) *Prevention of mental disorders, alcohol and other drug use in children and adolescents*. American Academy of Child Adolescent Psychiatry, Prevention monographs no. 3. Rockville, MD.

Reid, J. B., Patterson, G. R., & Snyder, J. (2002). (red.), *Antisocial behavior in children and adolescents: Developmental theories and models for intervention*. Washington, DC: American Psychological Association.

Reyno, S. M. & McGrath, P. J. (2006). Predictors of parent training efficacy for child externalizing behaviour problems – a meta-analytic review. *Journal of Child Psychology and Psychiatry*, 47, 99-111.

Robins, L. N. (1966). *Deviant children grown up: A sociological and psychiatric study of sociopathic personality*. Baltimore: Williams, & Wilkins. (her: refert fra Robins, 1978).

Robins, L. N. (1978). Sturdy childhood predictors of adult antisocial behavior: Replications from longitudinal studies. *Psychological Medicine*, 8, 611-622.

Rodick, J. Dj. & Henggeler, S. W. (1980). The short-term and long-term amelioration of academic and motivational deficiencies among low-achieving inner-city adolescents. *Child Development*, 51, 1126-1132.

Rogers, E. (1995). *Diffusion of innovation*. 4th edn. New York: Free Press.

Rosenzweig, S. (1936). Some implicit common factors in diverse methods of psychotherapy: At last the DoDo said: "Everybody has won and all must have prices". *American Journal of Orthopsychiatry*, 6, 412-415.

Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2004). *Evaluation. A systematic approach*. 7th edn. Thousand Oaks: Sage.

Rundskriv Q-12/97 (1997). Oslo: Barne- og familiedepartementet.

Rutter, M., Giller, H., & Hagell, A. (1998). *Antisocial behaviour by young people: The main messages from a major new review of the research*. London: Cambridge University Press.

Rønnestad, M.H. (2008). Evidensbasert praksis i psykologi. *Tidsskrift for Norsk Psykologforening*, 45, 444-454.

Sackett, D.L., Richardson, W.S., Rosenberg, & Haynes, R.B. (1997). *Evidence-based medicine: How to practice and teach EBM*. New York: Churchill Livingstone.

Sackett, D.L., Rosenberg, W.M., Gray, J.A., Haynes, R.B., & Richardson, W.S. (1996). Evidence based medicine: What it is and what it isn't. *British Medical Journal*, 312, 71-72.

Sánchez, V., Steckler, A., Nitirat, P., Hallfors, D., Cho, H., & Brodish, P. (2007). Fidelity of implementation in treatment effectiveness trial of Reconnecting Youth. *Health Education Research*, 22, 95-107.

Sanders, M. (1999). Triple P Positive Parenting Program: Toward an empirically validated multilevel parenting and family support strategy for the prevention of behavioral and emotional problems in children. *Clinical Child and Family Psychology Review*, 2, 71-90.

Schaeffer, C.M., & Borduin, C.M. (2005). Long-term follow-up to a randomized clinical trial of multisystemic therapy with serious and violent juvenile offenders. *Journal of Consulting and Clinical Psychology*, 73, 445-453.

Scherer, D. G., Brondino, M. J., Henggeler, S. W., Melton, G. B., & Hanley, J. H. (1994). Multisystemic family preservation therapy: Preliminary findings from a study of rural and minority serious adolescent offenders. *Journal of Emotional and Behavioral Disorders*, 2, 198-206.

Schulte, D., Kunzel, R., Pepping, G., & Schulte-Barenberg, T. (1992). Tailor-made versus standardized therapy for phobic patients. *Advances in Behaviour Research and Therapy*, 14, 67-92.

Schulz, K. F., Altman, D. G., Moher, D. (2010). CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trial. *BMC Medicine* 2010 8:18.

Serketich, W. J., & Dumas, J. E. (1996). The effectiveness of behavioral parent training to modify antisocial behavior in children: A meta-analysis. *Behavior Therapy*, 27, 171-186.

Selnes, B., & Larsen, E. (1974). *Fra avvik til ansvar*. Oslo, TANO forlag.

Sexton, T. L., & Alexander, J. F. (2002). Family-based empirically supported interventions. *The Counseling Psychologist*, 30, 238-261.

Sexton, T. L., Alexander, J. F., & Mease, A. L. (2004). Levels of evidence for the models and mechanisms of therapeutic change in couple and family therapy. I: Lambert, M.J. (red.), *Bergin and Garfield's handbook of psychotherapy and behavior change*. 5th edn. New York, Wiley.

Sexton, T. L., Ridley, C. R., & Kleiner, A. J. (2004). Beyond common factors: Multilevel-process models of therapeutic change in marriage and family therapy. *Journal of Marital and Family Therapy*, 30, 131-149.

Shadish, W. R., & Baldwin, S. A. (2002). Meta-analysis of MFT interventions. I: Sprenkle, D.H. (red.) *Effectiveness research in marriage and family therapy*. Alexandria VA: AAMFT.

Shadish, W. R., Montgomery, L., Wilson, P., Wilson, M., Bright, I., & Okwumabua, T. (1993). Effects of family and marital psychotherapies : A meta-analysis. *Journal of Consulting and Clinical Psychology*, 61, 992-1002.

Shirk, S. R. & Karver, M. (2003). Prediction of treatment outcome from relationship variables in child and adolescent therapy: A meta-analytic review. *Journal of Consulting and Clinical Psychology*, 71, 452-464.

- Skulberg, H. (2008). *Evidens og evidensdebattens betydning for utdanningssystemet*. Temanotat 6/2008. Oslo, Utdanningsforbundet.
- Smith, M. L., Glass, G. V., & Miller, T. I. (1980). *The Benefits of Psychotherapy*. Baltimore: John-Hopkins Press.
- Socialstyrelsen (2000). *Nationellt stöd för kunskapsutveckling inom socialtjänsten*. Stockholm: Socialstyrelsen.
- Solholm, R., Askeland, E., Christiansen, T., & Duckert, M. (2005). Parent Management Training - Oregon-modellen. Teori, behandlingsprogram og implementering i Norge. *Tidsskrift for Norsk Psykologforening*, 42, 587-597.
- Speer, D.C., & Greenbaum, P.E. (1995). Five methods for computing significant individual client change and improvement rates: Support for an individual growth curve approach. *Journal of Consulting and Clinical Psychology*, 63, 1044-1048.
- Spence, S. H. (1995). *Social Skills Training: Enhancing Social Competence with Children and Adolescents*. Windsor, Berkshire, UK: NFER-NELSON Publishing Company.
- Sprenkle, D. H., Blow, A. J., & Dickey, M. H. (1999). Common factors and other nontechnical variables in marriage and family therapy. I: Hubble, M.A., Duncan, B.L. & Miller, S.D. (red.). *The heart and soul of change: What works in therapy?*, Washington DC: American Psychological Association.
- Sprenkle, D.H., & Blow, A.J. (2004). Common factors and our sacred models. *Journal of Marital and Family Therapy*, 30, 113-129.
- Storvoll, E. (1997). *Barn og unge med alvorlige atferdsvansker: Hvem er de, og hvilken hjelp blir de tilbudt?* Rapport 21, Oslo: NOVA.
- Strother, K., Swenson, M. E., & Schoenwald, S. K. (1998). *Multisystemic therapy organizational manual*, Charleston, MST-services.
- Stubbs, J., Crosby, L., Forgatch, M. S., & Capaldi, D. M. (1998). *Family and peer process code: A synthesis of three Oregon Social Learning Center behavior codes* (Training manual.). Oregon Social Learning Center.
- Sundell, K., Hansson, K., Löfholm, C. A., Olsson, T., Gustle, L-H., & Kadesjö, C. (2006). *Multisystemisk terapi för ungdomar med allvarliga beteendeproblem. Resultat efter sex månader*. Redovisningar av forskningsresultat från IMS och EpC. Stockholm: Institut för utveckling av Metoder i Socialt arbete (IMS), Socialstyrelsen.

- Sundell, K., Hansson, K., Löfholm, C.A., Olsson, T., Gustle, L.H., & Kadesjö, C. (2008). The transportability of multisystemic therapy to Sweden: short-term results from a randomized trial of conduct-disordered youths. *Journal of Family Psychology*, 22, 550-560.
- Sørli, M.-A., Amlund-Hagen, K., & Ogden, T. (2008). Social competence and antisocial behavior: Continuity and distinctiveness across early adolescence. *Journal of Research on Adolescence*, 18, 121-144.
- Sørli, M.-A., Ogden, T., Solholm, R., & Olseth, A. R. (2010). Implementeringskvalitet – om å få tiltak til å virke: En oversikt. *Tidsskrift for Norsk Psykologforening*, 47, 315-321.
- Tallman, K., & Bohart, A. C. (1999). The client as a common factor: Clients as self healers. I: Hubble, M.A., Duncan, B.L., & Miller, S.D. (red.). *The heart and soul of change: What works in therapy?*, Washington DC: American Psychological Association.
- Tengvald, K. (2001). Kunskapsutveckling av socialt arbete – för klienters bästa, *Socionomen*, 1, 22-27.
- Trudeau, K. J., Mostofsky, E., Stuhr, J. K., & Davidson, K. W. (2008). Explanation of the CONSORT statement with application to psychosocial interventions. I: Nezu, A.M. & Nezu, C. M. (red.) *Evidence based outcome research. A practical guide to conducting randomized controlled trials for psychosocial interventions*. Oxford: Oxford University Press.
- Vedeler, G.H. (1973). *Samspill*. Oslo: Gyldendal forlag.
- Vinnerljung, B. Sallnäs, M., & Westermark, P. K. (2001). *Sammanbrott vid tonårsplaceringar – om ungdomar i fosterhem och på institution*. Rapport från CUS. Stockholm: Centrum för utvärdering av socialt arbete, Socialstyrelsen.
- Walker, H.M., Ramsey, E. & Gresham, F.M. (red.) (2004). *Antisocial behavior in school: Evidence-based practices*. Belmont, US.
- Waltz, J., Addis, M. E., Koerner, K., & Jacobson, N. S. (1993). Testing the integrity of psychotherapy protocol: Assessment of adherence and competence. *Journal of Consulting and Clinical Psychology*, 61, 620-630.
- Wampold, B.E. (2001). *The greath psychotherapy debate: Models, methods and findings*. Mahwah NJ: Erlbaum.
- Wampold, B. E., Mondin, G. W., Moody, M., Stich, F., Benson, K., & Ahn, H. (1997). A meta-analysis of outcome studies comparing bona fide psychotherapies: Empirically, "All must have prizes". *Psychological Bulletin*, 122, 203-215.

Webb, C. A., DeRubeis, R. J., & Barber, J. P. (2010). Therapist adherence/competence and treatment outcome: A meta-analytic review. *Journal of Consulting and Clinical Psychology*, 78, 200-211.

Webster-Stratton, C. H., & Herbert, M. (1994). *Troubled families - problem children: Working with parents: A collaborative process*. Chichester: John Wiley & Sons.

Weersing, V. R. & Weisz, J. R. (2002). Mechanisms of action in youth psychotherapy. *Journal of Child Psychology and Psychiatry*, 43, 3-29.

Weiss, C. (1979). The many meanings of research utilization. *Public Administration Review*, 39, 426-431.

Weiss, B. H., Catron, T., Harris, V., & Phung, T. M. (1999). The effectiveness of traditional child psychotherapy. *Journal of Consulting and Clinical Psychology*, 67, 82-94.

Weisz, J. R. (1997). Effects of interventions for child and adolescent psychological dysfunction: Relevance of context, developmental factors, and individual differences. I: Luthar, S. S., Burack, J. A., Cicchetti, D., & Weisz, J. R. (red.). *Developmental psychopathology. Perspectives on adjustment, risk and disorder*. Cambridge: Cambridge University Press.

Weisz, J. R. (2004). *Psychotherapy for children and adolescents: Evidence-based treatments and case examples*. Cambridge: Cambridge University Press.

Weisz, J. R., Donenberg, G. R., Han, S. S., & Weiss, B. (1995). Bridging the gap between lab and clinic in child and adolescent psychotherapy. *Journal of Consulting and Clinical Psychology*, 63, 688-701.

Weisz, J. R., Jensen-Doss, A., & Hawley, K. M. (2005). Youth therapy outcome research: A review of critique of the evidence base. *Annual review of psychology*, 56, 337-363.

Weisz, J. R., Jensen-Doss, A., & Hawley, K. M. (2006). Evidence-based youth psychotherapies versus clinical care. A meta-analysis of direct comparisons. *American Psychologist*, 61, 671-689.

Weisz, J. R., Sandler, I. N., Durlak, J. A., & Anton, B. S. (2005). Promoting and protecting youth mental health through evidence-based prevention and treatment. *American Psychologist*, 60, 628-648.

Weisz, J. R., Weersing, V. R., & Henggeler, S. W. (2005). Jousting with straw men: Comment on Westen, Novotny and Thompson-Brenner (2004). *Psychological Bulletin*, 131, 418-426.

Westen,D., Novotny, C.M., & Thompson-Brenner (2004). The empirical status of empirically supported psychotherapies: Assumptions, findings, and reporting in controlled clinical trials. *Psychological Bulletin*, 130, 631-663.

Westen, D., Novotny, C. M., & Thompson-Brenner, H. (2005). EBP does not equal EST: Reply to Crits-Christoph et al. (2005) and Weisz et al. (2005). *Psychological Bulletin*, 131, 427-431.

Wickstrøm, L., Skogen, K., & Øia, T. (1996). The increased rate of conduct problems in urban areas: what is the mechanism? *Journal of American Academy of Child and Adolescent Psychiatry*, 35, 471-479.

Winters, K. C., & Henly, G. (1989). *The personal experiences inventory*. Los Angeles: Western Psychological Services.

Wright, C. C., Sim, J. (2003). Intention-to-treat approach to data from randomized controlled trials: A sensitivity analysis. *Journal of Clinical Epidemiology*, 56, 833-842.

Zeiner, P., Backe-Hansen, E., Eskeland, S., Ogden, T., Rypdal, P., & Sommerschild, H (1998). *Barn og unge med alvorlige atferdsvansker. Hva kan nyere viten fortelle oss og hva slags hjelp trenger de?* Oslo, Norges Forskningsråd.

APPENDIKS I

Prinsipp 1. Den primære hensikten med utredningen er å forstå den gjensidige tilpasningen mellom de identifiserte problemene og deres bredere systemiske kontekst. Det betyr at problemene som ungdommen henvises for kartlegges og analyseres i forhold til kjennetegn ved familien, skolen, fritidsmiljøet og andre miljøfaktorer. Atferdsproblemene utredes i ungdommens sosiale nettverk med vekt på å forstå forhold som direkte eller indirekte bidrar til å vedlikeholde dem. Utredningen er en kontinuerlig prosess omtales som "the MST loop" der MST-terapeuten i samarbeid med familien (1) formulerer hypoteser om hvordan observerbare interaksjoner og atferd bidrar til ungdommens problemer, (2) samler data som kan bekrefte eller avkrefte hypotesene, (3) setter i verk intervensjoner, (4) observerer om intervensjonene fører til endringer i problematferden, (5) identifiserer eventuelle barrierer som hindrer et vellykket resultat og (6) iverksetter intervensjoner som kan nøytralisere disse hindringene. Terapeutens oppgave er å utrede de ulike deltakernes syn på sterke sider, problemer og behov for deretter å integrere de ulike synspunktene og forene divergerende perspektiver i en helhetlig forklaringsmodell. Arbeidshypotesen beskriver sammenhengen mellom problematferden og den systemiske konteksten, og hypotesen settes på praktisk prøve gjennom å endre de identifiserte påvirkningsfaktorene.

Prinsipp 2. Terapeutiske kontakter vektlegger det positive og bruker systemisk styrke for å fremme forandring. Systemisk styrke handler om å utnytte ressurser og beskyttende faktorer hos ungdommen selv og i hans eller hennes nære omgivelser for å redusere problematferd og fremme kompetanse. I all kommunikasjon med og om familien brukes et ikke-nedsettende språk, både i muntlig og skriftlig form. Dermed skapes en kommunikasjonsramme som legger hovedvekten på positive forventninger, familiens ressurser og andre beskyttende faktorer MST-terapeutene bruker "positiv reformulering" og er rause i bruk av oppmuntring og positiv bekreftelse av konkret innsats fra familien. Den positive fokuseringen skal likevel ha en forankring i sakens realiteter slik at en ikke skaper urealistiske forventninger til hva som kan oppnås.

Prinsipp 3. Intervensjoner skal fremme ansvarlig atferd og redusere uansvarlig atferd hos familiemedlemmene. Heller enn å fokusere på psykopatologi eller avvik, formidles det at målet er å redusere uansvarlig og øke ansvarlig atferd hos ungdom og foreldre. Foreldrenes ansvar er å gi sine barn omsorg og dekke andre grunnleggende behov, oppdra og veilede dem, samt hjelpe dem til å utvikle kompetanse. For ungdommen handler ansvarlig atferd om å overholde familiens og samfunnets regler, utføre plikter i huset, gå på skolen (eller jobb) og ikke skade andre. Sosial og håndfast bekreftelse som ros, tillatelse og privilegier brukes for å fremme ansvarlig atferd mens ekstra arbeidsoppgaver i huset, eller tap av privilegier brukes som negative konsekvenser.

Prinsipp 4. Intervensjonene er "her og nå" og handlingsfokusede, rettet mot konkrete og veldefinerte problemer. Dette prinsippet signaliserer at en er mindre opptatt av fortiden og prioriterer tiltak som forbedrer situasjonen i familien akkurat nå. Handlinger og hendelser som ligger fjernere i tid og rom prioriteres lavere enn de konkrete omstendighetene som umiddelbart og direkte påvirker familiemedlemmenes atferd. Oppgavene er handlingsorienterte i den forstand at de aktiviserer ungdommen og resten av familien i problemløsende arbeid. Det øker mulighetene for at de raskt opplever positive resultater og motiveres til ytterligere innsats for å nå de oppsatte målene. Veldefinerte mål vil i denne

sammenhengen si at de er objektive, målbare og frie for faguttrykk. Dermed er alle involverte klar over arbeidets retning og kriteriene som brukes for å måle resultatene. En kan kontrollere om tiltakene virker og avslutte behandlingen når målene er nådd.

Prinsipp 5. Intervensjoner rettes mot atferdssekvenser innenfor eller mellom multiple systemer. Arbeidet fokuserer på hvordan atferden i skolen påvirkes av hendelser i hjemmet og vice-versa, og hvordan atferden i skole og hjem preges av hendelser på fritiden. Multiple systemer brukes her for å beskrive de sosiale delsystemene som det sosiale nettverket består av, for eksempel skoleklassen, vennegruppen og familien. Arbeidet for å restrukturere ungdommens økologi retter seg mot familierelasjonene og familieinteraksjonen samt samhandlingen med det sosiale nettverket. Det arbeides med samhandlingsmønstre i familien, men også med forholdet mellom familien og deres sosiale omgivelser. En er særlig oppmerksom på hvordan atferden i én setting kan være funksjonelt knyttet til hendelser i andre settinger for eksempel hvordan atferd i skolen kan reflektere hendelser i familien eller blant venner. Det som særlig skiller MST fra andre tilnærminger er vekten som legges på samhandling på tvers av sosiale systemer, som for eksempel når foreldrene forsøker å få ungdommen ut av avvikende vennegrupper og inn i mer prososiale jevnaldningsmiljøer.

Prinsipp 6. Intervensjonene tilpasses utviklingstrinnet og ungdommens utviklingsmessige behov. Både ungdommen og familien gjennomgår utviklingsfaser og deres forutsetninger og behov endrer seg over tid. For de yngste ungdommene vil foreldrenes tilsyn og kontroll være viktige temaer, mens en hos de eldre ungdommene legger mer vekt på egen medvirkning og forberedelser til et uavhengig liv. De jevnaldrenes betydning varierer også med ungdommens alder. Regulering av kontakten med antisosiale venner er særlig viktig i fasen hvor ungdommen er særlig avhengig av dem og dermed lett påvirkbar. Men alder er ikke alltid en god indikator på kognitiv og sosial utvikling så oppgavene og forventningene må tilpasses ungdommens reelle utviklings- og funksjonsnivå.

Prinsipp 7. Intervensjonene bør kreve daglig eller ukentlig innsats av familiemedlemmene. Familien kan løse sine problemer raskere dersom alle familiemedlemmene og andre involverte sammen og samtidig setter kreftene inn på å nå behandlingsmålene. Daglig eller ukentlig innsats gjør det også lettere å evaluere og måle framgang og eventuelt korrigere tiltaksplanen.

Prinsipp 8. Intervensjonenes effektivitet evalueres kontinuerlig fra flere perspektiver. Dette gjør det mulig å overvåke hvilke endringer som finner sted i ungdommens økologi, og gir en umiddelbar tilbakemelding om hvor vellykket intervensjonene er. Det forventes raske og målbare resultater, som regel etter én til to uker. Hvis resultatene uteblir, revurderes analysen av problemet, arbeidshypotesen, eller alternative intervensjoner iverksettes. Hvis innsatsen ikke gir resultater forsøker en å unngå ”mer av det samme”. Heller enn å øke frekvensen eller intensiteten i eksisterende tiltak ser en etter alternative måter å angripe problemene på. Det er viktig at familien og andre får tilbakemelding fra noen som har observert eventuelle positive endringer i ungdommens atferd. Personer som kan gi slik tilbakemelding rekrutteres fra hele ungdommens nettverk. MST-terapeuten prioriterer informanter som har de beste forutsetningene og mulighetene for å observere atferden det gjelder.

Prinsipp 9. Intervensjonene bør utformes slik at de fremmer generalisering og langtids vedlikehold av terapeutisk forandring. Generalisering og varighet av atferdsendring er et sentralt tema i all behandling. Målet er ikke bare å oppnå umiddelbare positive endringer, men også sikre at disse vedvarer og gir utslag i andre sammenhenger og settinger enn der hvor de

ble oppnådd Derfor legges det vekt på å myndiggjøre foreldrene gjennom å: (1) utvikle selvhjelpsferdigheter som de kan bruke til å styrke sin sosiale økologi, (2) utvikle familiens kompetanse til å forhandle om nåværende og framtidige problemer og (3) utnytte beskyttende faktorer i det sosiale nettverket. Det problemløsende arbeidet ivaretas primært av foreldrene med terapeuten i en støttende og konsultativ rolle. Utvikling av myndiggjørende foreldrepraksis forutsetter at terapeuten ikke går for sterkt inn i familiesystemet og påvirker gjennom sin egen personlighet og innsats.

